

© 2013

KYONG EUN OH

ALL RIGHTS RESERVED

THE PROCESS OF ORGANIZING PERSONAL INFORMATION

By

KYONG EUN OH

A Dissertation submitted to the
Graduate School-New Brunswick
Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in

Communication, Information and Library Studies

Written under the direction of

Professor Nicholas J. Belkin

and approved by

New Brunswick, New Jersey

OCTOBER, 2013

UMI Number: 3606583

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3606583

Published by ProQuest LLC (2013). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

ABSTRACT OF THE DISSERTATION

THE PROCESS OF ORGANIZING PERSONAL INFORMATION

by KYONG EUN OH

Dissertation Director:

Professor Nicholas J. Belkin

The purpose of this study is to explore the process of organizing personal information in digital forms from a cognitive sociological perspective. This study also aims to develop a model that explains the personal information organization process.

Participants of this study were 18 academics in three different professional age groups. To collect data, a background questionnaire, a diary study, and two semi-structured interviews were conducted with each of the participants. After completing the background questionnaire, participants were asked to keep diary entries over a week whenever they decided to save or organize their personal information items in digital forms. Then, the first interview was conducted to ask how and why they organized their information in certain ways. About 2-4 weeks after the first interview, the second interview examined whether participants made any changes to the information items and categories discussed in the first interview. To analyze data, all the interviews were recorded, transcribed, and coded with a set of categories, which were developed based on the analysis of the literature. Then, the researcher confirmed, modified, or extended the

initial categories while analyzing the empirical data. Additionally, each of the diary entries then underwent content analysis.

The findings show that the process of organizing personal information consists of six stages: (1) Initiation; (2) Identification; (3) Temporary Categorization; (4) Examination/Comparison; (5) Selection/Modification/Creation; and (6) Categorization. These stages involve different actions, thoughts, and decisions, and various factors influence the process. The findings indicate that the organizing process is heavily influenced by participants' social environments. In this study, the personal information organization process (PIOP) model, which holistically shows the ranges of behaviors that are involved during the personal information organization process, has been developed.

This study deepens our understanding about people's information organization process by providing an integrated view of the process. In particular, by providing an explanation of the social aspects of personal information organization, this study makes a unique contribution to the field. This study also contributes to the development and design of various personal information devices, interfaces, and applications that support individuals' organization of their information.

ACKNOWLEDGEMENT

This dissertation would not have been possible without help and support of so many people. My deepest appreciation goes to my advisor, Dr. Nicholas Belkin. He generously shared his knowledge and time through weekly meetings for several years, and helped me learn and proceed step by step under his insightful guidance, warm encouragement, and incredible support. It was a great privilege to have an inspiring and caring advisor. I was also fortunate to have my dissertation committee members, who provided many interesting comments from different viewpoints: Dr. Nina Wacholder from information organization; Dr. Jacek Gwizdka from personal information management; and Dr. Eviatar Zerubavel from cognitive sociology. They helped me view my topic from various perspectives, which enriched my dissertation. I would also like to thank each of the participants of my dissertation. I could not have completed this research without their time and cooperation.

Over the past five years, I met many wonderful people at Rutgers. I would like to give my special thanks to Dr. Marilyn Tremaine. I learned a great deal while working with her, and always felt that I was supported both academically and emotionally. I also would like to thank to Dr. Marija Dalbello who was always encouraging and willing to help; Dr. Daniel O'Connor who made my Rutgers life so warm and joyful; Dr. Claire McInerney, Dr. Marie Radford, and Dr. Ya-Ling Lu for their caring advice on my academic career; and Ms. Joan Chabrak for her kindness and help. I also would like to acknowledge various research and teaching opportunities at Rutgers University, including my TAsip at the School of Communication and Information (SC&I) and RAship at the

Center for Advanced Information Processing (CAIP). I was not only supported financially, but also had great experiences.

I also thank my fellow doctoral students at the SC&I, especially to the 2008 cohorts and Korean doctoral students. Particularly, I want to thank to Sung-Un Kim who helped me from the beginning of my Ph.D. study, Eunkyung Lee who listened to my complaints with a warm smile, Sunny Lee who was always willing to share her experiences, and especially Goun Kim for her warmth and prayers. Very special thanks go to Soeyoon Choi. I cannot imagine my doctoral student life without the countless sighs and laughter I shared with her. They all became very precious memories of my life.

I also would like to acknowledge my professors at Yonsei University including Dr. Sang Wan Han, Dr. Tae Soo Kim, Dr. Young Mee Chung, Dr. Sung Been Moon, and Dr. Jee Yeon Lee. What I have learned at Yonsei University became a foundation of my knowledge in library and information science and rigorously trained me to develop capabilities in conducting research. I truly appreciate what they have taught me as well as their encouragement. Especially, I would like to thank my advisor, Dr. Giyeong Kim for his caring advice and continuing warm support.

I also thank to my old friends in Korea, especially Na Hyun Yoo, Kihyun Seung, Mi-young Kim, Eun Ji Jang, Jae Kyung Kim, and Miyoung Park for their constant friendship and emotional support from the opposite side of the earth. My special thanks go to Ju Young Kim who made me smile and laugh even during very rough times. I deeply appreciate your warmth, maturity, and understanding.

I would like to express my heartfelt gratitude to my family, one of the greatest and happiest blessings of my life. My warmest appreciation goes to my parents. I always felt

that I am deeply loved, cared, and supported by them. I also thank with special love to my two amazing sisters, Sang Eun Oh and Sangmee Oh who have always been my very best friends and great colleagues.

Finally, I give thanks to God. I am so grateful that my Father, who created every little part of me and knows me better than anyone else, chose library and information science as my academic field. He faithfully guided me and walked together with me as He always did, is, and will. I know that He always has better plans for me than mine, and I sincerely hope to live my life according to His will.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENT	iv
LIST OF TABLES	xii
LIST OF FIGURES	xv
CHAPTER 1 INTRODUCTION	1
1.1 Problem Statement.....	1
1.2 Purpose of the Study	4
CHAPTER 2 LITERATURE REVIEW	6
2.1 Categorization Theories	6
2.1.1 The Classical Theory of Categorization	6
2.1.2 Family Resemblances	8
2.1.3 Prototype Theory, Context Theory, and Fuzzy Set Theory	9
2.2 Information Behavior Studies	12
2.2.1 Information Organizing Behaviors	12
2.2.2 Processes of Information Behaviors	14
2.3 Personal Information Management.....	20
2.3.1 Personal Information Management Activities	20
2.3.2 Managing Personal Information in Different Forms	21
2.3.3 Organizing Personal Information.....	26
CHAPTER 3 THEORETICAL FRAMEWORK.....	37
3.1 Categorization in Cognitive Sociology.....	37

3.1.1 Lumping and Splitting	37
3.1.2 Thought Community	38
3.1.3 Typifications	40
3.1.4 Anomaly and Ambiguity.....	41
3.2 The Initial Model	43
3.2.1 Initiation	44
3.2.2 Identification	45
3.2.3 Examination	45
3.2.4 Comparison	46
3.2.5 Selection/Modification/Creation.....	46
3.2.6 Categorization	47
3.2.7 Underlying Factors.....	47
3.3 Research Objectives.....	48
CHAPTER 4 METHODS	50
4.1 Methods Used in Previous Work	50
4.1.1 Grand (Guided) Tour Interview	50
4.1.2 Think Aloud	51
4.1.3 Semi-structured Interview.....	51
4.1.4 Questionnaire	52
4.1.5 Video Recording	53
4.1.6 Diary Study	53
4.2 Sources of Data	56
4.2.1 Background Questionnaire.....	56

4.2.2 Diary Study	56
4.2.3 Interviews.....	57
4.3 Recruiting Participants.....	58
4.3.1 Recruiting Criteria	58
4.3.2 Recruiting Process	60
4.4 Data Collection	61
4.5 Data Analysis	63
4.6 Pilot Studies	63
4.6.1 First Pilot Study	63
4.6.2 Second Pilot Study.....	65
CHAPTER 5 FINDINGS.....	71
5.1 Characteristics of Participants.....	71
5.1.1 Demographic Information.....	72
5.1.2 Thought Communities	75
5.2 The Process of Organizing Personal Information.....	79
5.2.1 Stages	80
5.2.2 Behavioral Aspect.....	90
5.2.3 Cognitive Aspect.....	98
5.2.4 Decisions.....	110
5.2.5 Factors.....	124
5.2.6 Influence of Thought Communities	168
5.2.7 The PIOP (Personal Information Organization Process) Model.....	210
CHAPTER 6 DISCUSSION.....	215

6.1 The PIOP Model	215
6.1.1 Stages: Temporary Categorization and Types of Information Items	215
6.1.2 Actions: Reluctance to Change the Structure of Organization	217
6.1.3 Thoughts: Messiness, Typifications, and Category's Blurred Boundary	219
6.1.4 Decisions: Graded Structure of Categories	221
6.1.5 Factors: Comparative Analysis	222
6.1.6 Model: Dynamic and Holistic Process of the Information Organization	226
6.2 Social Aspect of the Process of Organizing Personal Information	230
6.2.1 The Influence of the Thought Community: Mind Structure	230
6.2.2 The Influence of Professional Age: Socialization	236
CHAPTER 7 CONCLUSION	240
7.1 Answering the Research Questions	240
7.2 Implications	242
7.2.1 Implications for Research	242
7.2.2 Implications for Practice	244
7.3 Limitations	245
7.4 Future Research	247
APPENDICES	249
Appendix 1: Recruitment Letter	249
Appendix 2: Informed Consent Form	251
Appendix 3: Audio/Videotape Addendum to Consent Form	254
Appendix 4: Initial Background Questionnaire	255
Appendix 5: Diary Instruction	256

Appendix 6: Diary Template	257
Appendix 7: Initial First Interview Questions	258
Appendix 8: Second Interview Questions.....	259
Appendix 9: Revised Background Questionnaire.....	260
Appendix 10: Revised First Interview Questions	261
Appendix 11: Part of the Coding Scheme.....	262
Appendix 12: Codebook	263
BIBLIOGRAPHY.....	275
CURRICULUM VITAE.....	283

LIST OF TABLES

Table 1 Organizing Strategies of Different Forms of Personal Information	32
Table 2 Types of Data, and Advantages and Disadvantages of Using Each Method.....	54
Table 3 Participants in Each Sub-group.....	71
Table 4 Participants' ID	72
Table 5 Participants' Gender in Each Sub-group	73
Table 6 Participants' Age Group	73
Table 7 Participants' Ethnicity.....	74
Table 8 Participants' Department	74
Table 9 Participants' Professional Age.....	75
Table 10 Number of Diary Entries.....	80
Table 11 Changes to the Stages of the Process.....	87
Table 12 Codes for Behavioral Aspect of the Process.....	91
Table 13 Codes for Cognitive Aspect of the Process	98
Table 14 Codes for Decisions Made During the Process	111
Table 15 Codes for Factors Impact on the Process.....	125
Table 16 Frequency of Occurrences of the Factors Impact on the Process	167
Table 17 Ubiquity of Factors on the Process	168
Table 18 Frequency of Occurrences of the Factors Impact on the (2) Identification Stage	170
Table 19 Frequency of Occurrences of the Factors Impact on the (4) Examination/Comparison Stage and the (5) Selection/Modification/Creation Stage.....	173

Table 20 Categories Used By Multiple Participants.....	178
Table 21 Stages of the Process of Organizing Academic and Non-academic Information Items.....	182
Table 22 Actions Taken During the Process of Organizing Academic and Non-academic Information Items.....	183
Table 23 Thoughts Shown During the Process of Organizing Academic and Non- academic Information Items	185
Table 24 Decisions Made During the Process of Organizing Academic and Non- academic Information Items	187
Table 25 Top Five Factors Influencing the Process of Organizing Academic and Non- academic Information Items	189
Table 26 Average Professional Age of Participants in Different Sub-groups	192
Table 27 Number of Academic and Non-academic Information Items of Participants in Different Sub-groups	193
Table 28 Categories Used by Multiple Participants in Different Sub-groups	195
Table 29 Stages of the Information Organization Process of Participants in Different Sub- groups.....	197
Table 30 Actions Taken During the Information Organization Process of Participants in Different Sub-groups	198
Table 31 Thoughts Shown During the Information Organization Process of Participants in Different Sub-groups	200
Table 32 Decisions Made During the Information Organization Process of Participants in Different Sub-groups	202

Table 33 Top Five Factors Impact on the Information Organization Process of Participants in Different Sub-groups	204
Table 34 Top Ten Factors Impact on Personal Information Organization	222
Table 35 Top Ten Factors Impact on Personal Information Organization Grouped into Categories	226

LIST OF FIGURES

Figure 1 The Model of the Information Process (Cole, 1997).....	15
Figure 2 Information Search Process (Kuhlthau, 1993 [1983]).....	16
Figure 3 Document Selection Model (Wang & Soergel, 1998).....	19
Figure 4 Percentage of Files by Modification Date (Gonçalves & Jorge, 2003).....	29
Figure 5 The Stance of Cognitive Sociology	37
Figure 6 The Initial Model of the Personal Information Organizing Process (PIOP)	44
Figure 7 Data Collection Process	62
Figure 8 Stages of the Personal Information Organization Process in the Initial Model ..	81
Figure 9 Stages of the Personal Information Organization in the Final Model	90
Figure 10 Behavioral Aspect of the Personal Information Organization Process	97
Figure 11 Cognitive Aspect of the Personal Information Organization Process	110
Figure 12 Decisions Made During the Personal Information Organization Process	124
Figure 13 The Depth of the Folder Structure for Academic and Non-academic Information Items.....	191
Figure 14 An Example of the PIOP Model for an Information Item.....	211
Figure 15 An Example of the PIOP Model for an Information Item with Future Decisions	212
Figure 16 The Personal Information Organizing Process (PIOP) Model	213

CHAPTER 1 INTRODUCTION

1.1 Problem Statement

Personal information is the information one keeps for personal use either directly or indirectly. Personal information management are the activities a person performs in order to create, find, organize, keep, maintain, use and distribute the information needed to meet life's various goals. In recent years, interest in personal information management has been increasing both as a technology topic and a serious area of research, not only in library and information science but also in many other disciplines including cognitive psychology, database management, artificial intelligence, knowledge management and human-computer interaction (Jones, 2007; Teevan, Jones, & Bederson, 2006). The ultimate goal of personal information management is to support the activities a person performs to organize his/her daily life by managing personal information (Teevan et al., 2006). Managing personal information effectively is important as it is directly related to an individual and his/her daily life in pursuing various goals, performing diverse roles and responsibilities.

Effective personal information organization has various advantages. For instance, it makes it easy to find needed information (Jones, Phuwanartnurak, Gill, & Bruce, 2005; Taylor & Joudrey, 2008) which saves an individual's time and effort. In addition, effective organization of personal information successfully reminds and manages tasks, supports further understanding and learning about information, enhances confidence, makes people feel satisfied about having things in order, and even shapes the social impressions of one as a well-organized and productive person (Barreau & Nardi, 1995;

Cutrell, Dumais, & Teevan, 2006; Ducheneaut & Bellotti, 2001; Erickson, 2006; Jones, 2007; Jones, Phuwanartnurak, Gill, & Bruce, 2005; Malone, 1983; Marshall & Jones, 2006; Teevan et al., 2006; Whittaker & Sidner, 1996). In addition, organizing personal information is closely related to individuals' personal productivity, satisfaction and creativity, which is connected to the productivity of their social or private groups (Barreau, 1995). Thus, organization of personal information plays a pivotal role in personal information management as well as individuals' lives by allowing an individual to make better use of his/her resources, including time, money, energy and attention, thereby improving the quality of his/her life (Jones, 2007).

However, despite the importance of personal information organization in our lives, many people have difficulty in organizing their personal information effectively (Bellotti et al., 2005; Lansdale, 1988; Ravasio, Schär, & Krueger, 2004; Whittaker & Sidner, 1996). In fact, when organizing personal information, people have to make various decisions. For instance, people have to decide (1) whether to include the information item in a certain category or not, (2) whether to create a new category or not, (3) whether to subdivide categories or not, (4) whether to create a superordinate category or not, and (5) whether to delete a certain category or not. Making such decisions is challenging mainly because information items can be categorized into several overlapping and fuzzy categories; they cannot be categorized into a neat categorization structure (Lansdale, 1988; Malone, 1983; Whittaker & Sidner, 1996). Because an information item has various aspects, it can be categorized into multiple categories. This means that, when an information item is categorized into a certain category, it is relevant to that category only in certain aspects (Lansdale, 1988).

Another reason that makes organizing personal information challenging is the fact that both information items and categories are not static but changing over time (Gottlieb & Dilevko, 2003; Jones, 2007; Jones & Teevan, 2007; Whittaker & Hirschberg, 2001). To be more specific, the value and usefulness of information items change as a person's interests and needs shift. Information that was extremely useful can become useless, or vice versa, in response to the changing interests and needs of a person. In a similar vein, information that was relevant to certain task may become irrelevant or more relevant to other tasks. What makes personal information organization even more challenging is the fact that the future interests and needs of a person are often unpredictable (Bruce, 2005; Whittaker & Hirschberg, 2001).

In addition, information overload, which occurs when people receive more information than they can process, is making personal information management more difficult (Ducheneaut & Bellotti, 2001; Mackay, 1988). Moreover, development of various new technologies has caused information fragmentation such that personal information is scattered across devices and tools in various forms (Jones, 2004; Jones, 2008), which makes personal information management even more challenging. Research on personal information organization indicated that people spend a lot of time and effort on organizing personal information, yet often find it difficult and unsuccessful (Bellotti et al., 2005; Lansdale, 1988; Malone, 1983; Ravasio, Schär, & Krueger, 2004).

However, despite its importance and close relationship with our lives, little is known about the process of organizing personal information. To be more specific, the mental process of organizing personal information has not been thoroughly explored yet. In addition, there appears to have been no research that has specifically examined

decisions that are made during the process of organizing personal information. Moreover, previous studies have rarely taken society into account, even though social environments heavily influence and constrain the personal information organizing process (Brekhus, 2007; Zerubavel, 1997).

Thus, it is important to explore the process of organizing personal information from a cognitive sociological perspective to advance our knowledge and understanding about people's information organization process, and to develop strategies and tools that effectively support personal information management so that people can make the most of their personal information, time and abilities in their everyday life.

1.2 Purpose of the Study

The proposed research aims to explore the process of organizing personal information. In particular, this study aims to identify different stages of the process as well as investigate actions shown during the process, thoughts involved during the process, primary decisions made during the process, and various factors that impact on the process of organizing personal information. This study will present a model which shows ranges of behaviors involved during the whole process of organizing personal information. It is important to note that this study takes the social foundation of individuals into account, as it assumes that information organization is a process of construction, and society constructs people's cognition and how they organize information. This study views individuals' classification systems and use of metaphors in classifying objects as the products of socialization into social rules of relevance and

irrelevance (Brekhus, 2007). Thus, the whole process of organizing personal information is examined and understood in this framework.

To limit the scope of the study, the form of personal information examined is restricted to electronic (i.e. *digital*) forms that are saved in personal devices including personal computers and mobile devices. Personal information in digital forms includes text files, image files, music files and other multi-media files. Among various forms of personal information, the electronic form was selected because it has become one of the most widely used forms of personal information, and is still increasing in its significance. In addition, although organization of information includes both establishing and populating categories, and establishing various relationships among those categories, this research study primarily focused on the establishment of, and assignment to categories, and to some extent the establishment of only hierarchical relationships among them.

CHAPTER 2 LITERATURE REVIEW

2.1 Categorization Theories

Categorization is a key process in organization of knowledge. Categorization has been studied and discussed for a long time by philosophers, psychologists, anthropologists, sociologists, linguists, and information scientists from different perspectives with different focuses. Especially, philosophers and psychologists made a great effort to categorize knowledge for many centuries (Taylor & Joudrey, 2008). While their focuses were specifically on neither categorization of personal information nor people's classification behaviors, these theories provide good background knowledge to understand the structures and the characteristics of a category.

2.1.1 The Classical Theory of Categorization

The discussions of categorization date back to ancient Greece when Aristotle (2007 [B.C.350]) stated that every object of human apprehension can be categorized under ten categories: substance; quantity; quality; relation; place; time; position; condition; action; and passion. In his view, every object has defining properties which solely determine whether that object can be classified into a certain category or not (Iyer, 1995; Lakoff, 1987). For instance, to be categorized as a rectangle, an object must be a closed figure that has four sides with four interior angles of 90°. If any of these is false, then it is not a rectangle (Iyer, 1995). Thus, according to Aristotle's theory, which is known as 'the classical theory of categorization', categories have clear and solid boundaries that are determined by defining properties. While this view ostensibly seems

reasonable as people do classify objects based on certain shared properties of objects, this theory fails to provide explanations for various challenges people face when organizing objects. The classical theory of categorization presumes that categorization can be done perfectly with a distinct boundary based on two assumptions. First, all objects in a category are equivalent in that the properties of the category are equally shared by all members in the category. In this perspective, one object in a category cannot be more typical than other objects in the category. For example, in the case of 'rectangle' category, which includes any object that has a closed figure that has four sides with four interior angles of 90° , every member of the category will be the same in terms of typicality. Second, categories are independent of a person who is categorizing because categories are defined solely by defining properties inherent in the objects (Lakoff, 1987). Thus, when applied to personal information organization, this theory is unsound for two reasons. First, most personal information items in a category do not have concrete and static defining properties that all of the objects in the category share. Second, in personal information organization, categorization often does not make any sense without understanding and reflecting the specific needs and interests of the person who is categorizing his/her own information items. For instance, if a person created a quick category for information items he/she wants to print out, such as a map of the street, a thank you note, and assignments for his/her class, it will be almost impossible for another to define equally shared properties that are inherent in the information items that determine their membership of the category. In addition, this category would not make any sense without taking that person into account. Barsalou (1983) also stated that in reality, people construct a number of 'ad hoc categories', which are categories that are

“created spontaneously for use in specialized contexts” (p. 211). However, while the classical theory obviously fails to provide a framework for studying the process of organizing personal information, it is useful in a sense that by showing how it cannot explain the process of personal information organization, it exposed that two assumptions of the classical theory of categorization are wrong. As a matter of fact, although classical theory dominated people’s understanding of categorization over the centuries, a number of researchers have challenged this view. They showed how it fails to explain the categorization process and suggested new theories that provided a sound framework for studying personal information organization process.

2.1.2 Family Resemblances

Wittgenstein was a philosopher who challenged the classical theory of categorization. One of the main points made by Wittgenstein (1953) was that there is no clear boundary to a category or concrete defining properties that determine the membership of the category. Rather, objects are grouped together by “family resemblances” (p.32) just like members of a family who share some common properties like color of hair or eyes, but do not have defining properties that determine them as members of a family. Later, many other theories and studies were proposed that supported Wittgenstein’s view. For instance, Lounsbury (1956), who studied the Native American kinship system, found that categories that seem definite in one culture are not the same categories in another culture. Although the main focus of the study was not on the categorization of information objects, Wittgenstein’s family resemblances provide a good explanation for why information items can be categorized into several overlapping

categories. An information item has a variety of aspects, and it can be categorized into multiple categories based on each of various aspects that are similar to certain aspects of other information items. For instance, suppose a person created a PowerPoint slide that he/she is going to use in the class he/she is teaching in this semester. Based on which aspect is accounted for while categorizing, this file can be grouped with information items that are for teaching (purpose), that are for this semester (time), that are created by him/her (source) or that are in PowerPoint format (form). Thus, Wittgenstein's family resemblance explains why categorization cannot be done neatly and completely.

2.1.3 Prototype Theory, Context Theory, and Fuzzy Set Theory

Another approach that greatly challenged the classical theory of categorization and provided profound insights into our understanding of categories is Rosch and her associates' psychological research. Based on an array of empirical studies, Rosch and her associates introduced the concept of 'prototype' (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Rosch (1978) defined 'prototypes' of categories as "the clearest cases of category membership defined operationally by people's judgments of goodness of membership in the category" (p. 36). For instance, in most cases, 'rose' is a better example than 'portulaca grandiflora' for a 'flower' category. The central idea of the prototype theory is that (1) category structure is asymmetric and graded, which means that some members of a category are better examples than other members in a category, (2) the boundary of the category is not solid but blurry, and (3) the definition of categories, and the ranking of typicality of members depends highly on who is categorizing (Barsalou, 1983; Iyer, 1995; Rosch, 1978). The existence of the

graded structure and blurry boundary of categories has been studied and proved by many other researchers in different fields both theoretically and empirically. For instance, Berlin and Kay's (1969) study on color categorization found that the 'foci colors', which are the best example of the category, are similar across 20 unrelated languages (p.10). When the researchers asked participants from different countries to point to certain colors from a standardized chart of 329 color chips, participants chose all different shades of the color. However, when they asked participants to pick the best example of a certain color, most participants picked the same color (Berlin & Kay, 1969; Iyer, 1995). Thus, Berlin and Kay's study showed that there are better examples of a category and the structure of categories is graded. Ekman, Friesen and Ellsworth (1972), who studied facial gestures that express emotion across cultures, also found that there are seven basic emotions that have prototype status.

Another theory that is related to prototype theory is context theory (Medin & Schaffer, 1978). The general idea of a context theory of classification is that people organize their personal information objects by comparing new information objects with their previous information objects. Thus, the classification judgment derives from stored information (p. 207). In this perspective, prototype theory is an extension of context theory of classification (Barreau, 1995, p. 329). To be more specific, based on prototype theory, people organize information objects by using a prototype, which is an exemplar of the category that provides the basis for classification judgments (Medin & Schaffer, 1978, p. 216).

Zadeh (1965), a mathematician and electrical engineer, devised the famous 'fuzzy set theory', which showed the graded membership of a category. In his theory, he defined

‘fuzzy set’ as a class with a continuum of grades of membership (p. 339). His argument is that in reality, each object in a set has a value between 0 and 1, rather than either 0 or 1 (p. 338). By using mathematical operations including union, intersection, complement and relation, Zadeh showed that objects do not have precisely defined criteria for membership, but are on a continuum of grades of membership.

These theories are critical when studying the process of organizing personal information. By showing that categories have graded structures, it provides an explanation for why categorization of certain information items is easier or more difficult than for other information items. In particular, Rosch, Simpson and Miller (1976) found that when participants are asked to make judgments about whether a certain object is a member of a certain category or not, the judgment response was much faster for the objects that have been rated more prototypical (p. 498). Based on prototype theory, we can hypothesize that a person will find certain information items easier to categorize than other information items, and an information item that is easier to categorize will be a more typical member of the category than other information items. In addition, based on the results of the study by Rosch et al (1976), which reported that participants list the most prototypical objects first and most frequently when asked to recall members of the category (p. 498), we can hypothesize that when people examine previous categories they created while making categorization decisions for a new information item, it is very likely that they will first recall prototype information items rather than other information items in existing categories. Most importantly, these theories help understand why categorization cannot be done simply and perfectly. Especially, these theories provide a framework to view categorization in a natural way, and recognize that the existence of

imprecise categorization originates from “the absence of sharply defined criteria of class membership rather than the presence of random variables” (Zadeh, 1965, p. 339).

The studies introduced here all examined categorization, and provide good background knowledge in studying categorization. In particular, they help to understand the structure of a category, i.e., graded structure and blurred boundary. However, most of the studies examined categorization either conceptually, or investigated classification of objects in the natural world. Few have examined classification of information objects. In addition, the primary focuses of the study was not on the classification behavior of people.

2.2 Information Behavior Studies

One of the main research areas in information science is human information behavior, which studies “those activities a person may engage in when identifying his/her own needs for information, searching for such information in any way, and using or transferring that information” (Wilson, 1999, p. 249). While most information behavior studies have focused on information seeking behavior, there are few studies which examined information organizing behavior. In addition, previous studies that investigated the processes of information behaviors provide insights for this study.

2.2.1 Information Organizing Behavior

In the case of information behavior studies, while most studies focused on information seeking behavior, some studies examined how people classify information objects. Cooper’s (2004) study, which investigated the influence of society on

information organization, is one of the studies that is closely related to this research study. To examine how people's understandings of categories are influenced by society, Cooper carried out an interesting research study with 516 children in kindergarten to 4th grade. First, Cooper asked participants in each grade level to (1) suggest the information they thought most important to have in a library. Then, the terms most frequently named by participants in each grade level were selected as sort terms for that grade level. (2) Then the researcher showed each of 20 terms to the participants and asked whether they would put that 'book' on a shelf with the first one or on a separate shelf. After grouping them, the participants were asked to give each shelf a name. (3) Then, the researcher asked participants to form a group consisting of four children, and sort 58 terms and label each category without her assistance. (4) Last, the researcher asked participants in groups to sort terms, which were names of animals (p. 309-311). Collected data were entered into the spreadsheet and analyzed by generating a co-occurrence matrix, hierarchical clustering, and multidimensional scaling output for each grade level by using UCINET (p. 313). As a result, Cooper found that society influences information organizing behavior so that children move from their personal categories for information toward more non-personal categories for information to interact with the library as they get older (p. 299). Cooper's study is relevant to this research study not only because she examined information organizing behavior, but also because she explored the influence of society on information organizing behavior. However, Cooper's study dealt with organization of public information rather than personal information. Also, she examined the influence of primary socialization on information organizing behavior, while this research study

focuses on the impact of secondary socialization on the process of organizing personal information.

2.2.2 Processes of Information Behaviors

In information behavior research, there were several studies that investigated the *processes* of certain information behaviors such as the process of information seeking and selecting. They are helpful in the sense that they show how the processes of information behaviors have been studied. Cole (1997) conducted interviews with 45 history doctoral students and reported a pattern of cognitive activity that occurs while searching for information. Then he proposed a model, which he called ‘the model of the information process’ that is made up of five stages (p. 59). These stages are (1) opening of the information process, which occurs when a person has a difficulty in recognizing certain information because it is something new or unexpected, (2) representational cognitive activity, which is a person’s effort to relate new or unexpected information with his/her previous knowledge, (3) corroborating evidence looked for and found, which is a person’s seeking out some corroborating evidence that may support understanding new information, (4) closing of the information process, which occurs when a person successfully contextualizes new information by relating it with his/her previous knowledge, and (5) the effect of the information process, which is a modification of knowledge structure so that a person can perceive information in material which he/she could not understand previously (p. 60-63). This study is relevant to the proposed research as Cole identified different stages of information processing while focusing on

the cognitive aspect of the process. However, Cole did not take social influence into account. Cole's model of the information process is presented in Figure 1.

A five stage model of the information process for the 45 Ph.D. history students interviewed for the study	
Stage 1.	Opening of information process — conscious — unconscious
Stage 2.	Representational (cognitive) activity — set of alternative possibilities — conscious — unconscious
Stage 3.	Corroborating evidence sought and found (the "what?") — conscious — focused information seeking behaviour — unconscious — unfocused information seeking behaviour
Stage 4.	Closing of process (the "why?") — contextualization of the opening and corroborating data elements — internally located — externally located — focused information seeking behaviour
Stage 5.	Effect of process: Knowledge Structure is modified — the ability to "see" newness in material read before — focused information seeking to confirm "information"

Figure 1. The Model of the Information Process (Cole, 1997)

Kuhlthau (2005) also investigated the process of information search and developed a model, which represents common patterns people engage in during the process of information seeking (p. 230). This model is made up of six stages: (1) initiation, which occurs when a person recognizes a lack of knowledge or uncertainty; (2) selection, in which a person identifies a general area or problem; (3) exploration, which involves a person's encountering inconsistent information and increased uncertainty; (4) formulation, which is a person's formulating a focused perspective which diminishes

uncertainty; (5) collection, which is a gathering of relevant information; and (6) presentation, which occurs when the search is completed (p. 230-231). Kuhlthau investigated thoughts, feelings, and actions that are involved in each stage while developing her model, and asserted that “people experience the information search process holistically, with an interplay of thoughts, feelings, and actions” (p. 230). This view is closely related to this research study, which aims to provide a holistic view of the organizing process. However, it is different from Kuhlthau’s approach in that while Kuhlthau examined the searching process of public information, this research explores the organizing process of personal information. In addition, while Kuhlthau did not take social influence into account, the proposed study uses cognitive sociology as a main framework that influences the whole process. Kuhlthau’s Information Search Process model is presented in Figure 2.

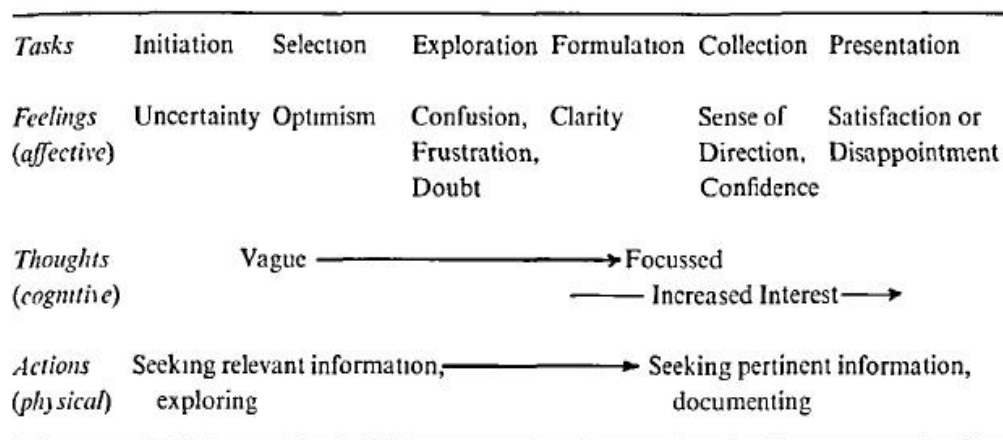


Figure 2. Information Search Process (Kuhlthau, 1993 [1983])

Chen and Dhar (1991) also investigated the cognitive processes that are involved during information retrieval on the web. Using think-aloud protocols, conducting interviews, recording the interactions, and using questionnaires, the researchers analyzed

the interactions between 34 users and a retrieval system both with and without the assistance of reference librarians (p.407). The researchers identified five distinctive process models of online information retrieval search strategies (p. 410). Those five models are (1) known-item-instantiation, (2) search-option-heuristics, (3) thesaurus-browsing, (4) screen-browsing, and (5) trial-and-error (p. 411-415). However, while this study identified process models, it mainly focused on the activities people perform while searching for information that is available via a library, rather than organizing personal information. In addition, they did not examine either the cognitive or social aspects of the process.

Wang and Soergel (1998) conducted an interesting study in which the authors examined how users select documents for their projects, with a special focus on the processes of document selection (p. 116). This study examined both components that are relevant to the users' decisions and cognitive processes that have occurred during the document selection (p. 116). Especially, the authors proposed a cognitive model of decision-making in the document selection process. To conduct this research, the researchers asked 25 participants with specific information needs to select relevant documents from a search results list, which the researchers obtained by searching databases on the participants' topics. The participants selected documents while using a think aloud protocol, and those verbal reports were analyzed. Then, the authors proposed a model of document selection. In this model, first, (1) a document is represented by a set of document information elements (DIEs) such as title, abstract, author, document type, language, etc. (p.117). (2) Then, these DIEs are processed to judge the documents based on eleven criteria from the content analysis. These include topicality, orientation/level,

discipline, novelty, quality, recentness, reading time, availability, special requisite, authority, and relation/origin (p. 122). (3) Then, these criteria are applied to judge document values (p.117). The authors introduced five values: epistemic value; functional value; conditional value; social value; and emotional value (p. 121). (4) Finally, people made the document selection decisions based on these values (p. 117). This whole process is governed by decision rules, which are elimination rule, multiple-criteria rule, dominance rule, scarcity rule, satisfaction rule, and chain rule (p. 128). In addition, personal knowledge plays a pivotal role in this process (p. 128). From content analysis, the authors identified 17 DIES from the data, and found that title was the most frequently used (57%). Among 11 criteria, topicality (65.3%), orientation/level (9.4%), and quality (9.4%) were frequently used. Among five values, epistemic value, functional value, and social value were salient (p. 128). This study is relevant to this research as the authors investigated the process together with decisions that are made during the process. However, Wang and Soergel examined the process of document selection rather than the personal information organizing process. In addition, the researchers did not take social influence into account. Wang and Soergel's document selection model is presented in Figure 3.

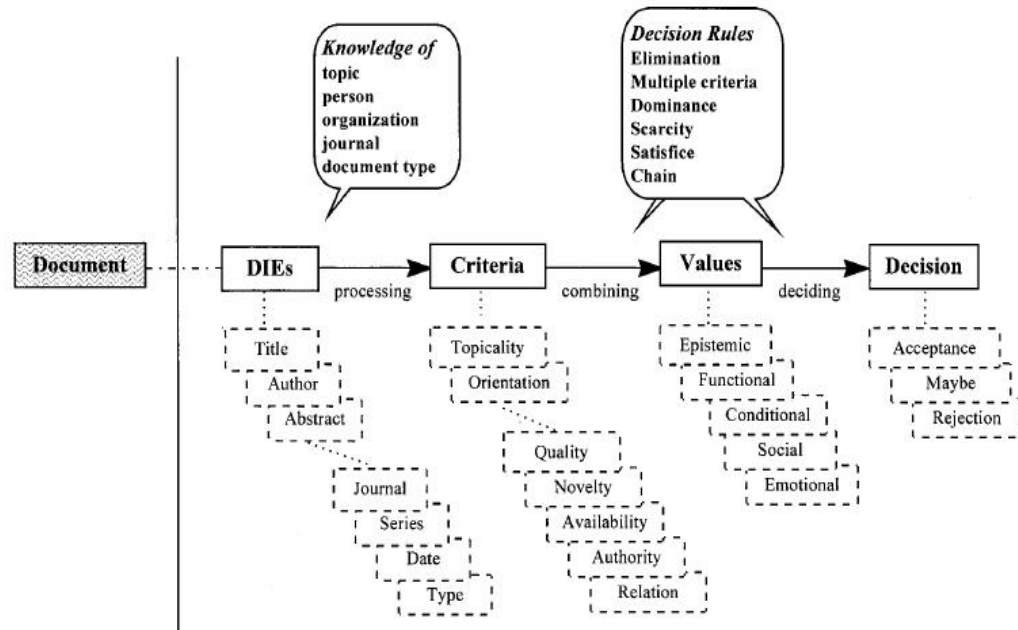


Figure 3. Document Selection Model (Wang & Soergel, 1998)

Thus, there are several studies that examined information organizing behavior or explored the process of certain information behaviors. In the case of studies that examined information organizing behavior, they help understanding on how to study classification behavior. However, these studies focused on organization of public information rather than personal information. Studies that investigated the processes of certain information behaviors are helpful in the sense that they show how processes of information behavior have been studied. Kuhlthau's study, which identified three realms of experience, which are affective (feelings), cognitive (thoughts) and physical (actions) along with the process, was particularly helpful in developing the starting model for this study. However, as previously mentioned, the main focus of these studies was not on the organization of information. More importantly, these studies did not relate society with information behavior.

2.3 Personal Information Management

In the information science field, as previously mentioned, there have been a number of insightful theories and studies that deal with how people seek new information (Bates, 1989; Belkin, 1980; Ellis 1989; Kuhlthau, 1991; Wilson, 1997). However, fewer studies focused on what happens after a person obtains new information as a result of information seeking (Whittaker, 2011). As a matter of a fact, in our everyday life, as much as people seek new information, people extensively engage in managing already found information, i.e., personal information management. Personal information management is “the practice and the study of the activities a person performs in order to acquire or create, store, organize, maintain, retrieve, use and distribute the information needed to complete tasks and fulfill various roles and responsibilities” (Jones, 2007, p. 453). Thus, personal information management covers various topics which examine a variety of aspects of how people deal with their personal information (Jones, 2007; Jones, 2008). As this study aims to explore the organizing process of personal information, it is important to understand the personal information literature and examine what has been studied and has not been investigated.

2.3.1 Personal Information Management Activities

The process of personal information management involves several different activities. Once a person acquires a certain information item by finding, receiving or creating information, he/she may keep it in his/her personal devices, organize it by grouping it with related information items and distinguishing it from other information items, maintain it, and re-find it when he/she needs it. These activities can be grouped

into three essential activities, which are: (1) keeping activities; (2) meta-level activities (organizing and maintaining); and (3) finding/re-finding activities (Jones, 2007; Jones & Teevan, 2007). These three primary activities are distinct yet closely related to each other (Jones, 2008; Whittaker, 2011). Keeping activities are the prerequisite activities for other personal information management activities, as personal information items can be organized and maintained when they are saved in certain locations. In addition, keeping activities can make both meta-level activities and finding/re-finding activities difficult, because the more a person keeps information items, the more he/she needs to organize and maintain them. Also, he/she needs to find information items among an ever-growing number of information items. Meta-level activities can facilitate or hinder finding/re-finding activities. When information items are organized and maintained in a way that is readily accessible, a person may find information items easily. However, when it is organized and maintained in a way that confuses a person, he/she will have difficulties in finding information items. How information items are found and re-found will also have an impact on how information items will be organized and maintained in the future. Thus, in personal information management, each activity influences the other activities.

2.3.2. Managing Personal Information in Different Forms

Most research in personal information management has centered on three essential activities of personal information management, which are finding/re-finding, keeping, and organizing personal information in a specific form. Primary forms of information items that have been the main focus of personal information management studies are paper-based information, emails, web information, and electronic files. Some

major issues that have been addressed in personal information management research are detailed below.

2.3.2.1 Paper-based Information

Early works in personal information management mostly investigated how people manage their paper-based information items such as books and paper-based documents, especially in their offices (Case, 1986; Cole, 1982; Kwasnik, 1989; Malone, 1983; Soper, 1976). Among these works, Malone's study identified two different strategies in managing personal information: piling and filing. Kwasnik's study investigated various factors that are associated with organizing personal information. These are two of the pioneering works that largely influenced later personal information research. General findings from other studies on the management of paper-based information items noted that people keep a large amount of paper-based information items (Case, 1986; Soper, 1976; Whittaker & Hirschberg, 2001), and have difficulties in managing paper-based information items. These difficulties include general difficulties such as requiring too much effort to organize information items in a way that facilitates future access (Case, 1991; Whittaker & Hirschberg, 2001), and unique difficulties such as spatial constraints in keeping and organizing information (Case, 1991). While personal information in digital form is growing, people still manage a large amount of paper-based information (Bondarenko, Janssen, & Driessen, 2010; Whittaker & Hirschberg, 2001). Thus, understanding how people manage paper-based information is still crucial in understanding various personal information management behaviors.

2.3.2.2 Email

Email management has been one of the main research areas in personal information management. A number of studies identified different strategies used in keeping and organizing emails (Bälter, 1997; Boardman & Sasse, 2004; Fisher, Brukh, Gleave, & Smith, 2006; Gwizdka, 2004; Mackay, 1988; Whittaker & Sidner, 1996). In addition, unique functions of email in managing personal information have been discussed by many researchers. The researchers found that people use email not only to communicate with others in remote places, but also to keep information (Jones, 2008; Whittaker, Bellotti, & Gwizdka, 2006; Whittaker & Sidner, 1996), to manage and remind tasks (Jones, 2008; Mackay, 1988; Whittaker, Bellotti, & Gwizdka, 2006; Whittaker & Sidner, 1996), and to manage documents, schedules, and contact information (Ducheneaut & Bellotti, 2001; Jones, 2008; Mackay, 1988; Whittaker, Bellotti, et al., 2006; Whittaker et al., 2004). The challenges involved in email management, which mostly originate from the unique characteristics of email, have also been discussed. The researchers stated that individual emails are difficult to organize and process because they contain diverse types of information such as tasks, documents, announcements, and schedules (Whittaker, Bellotti, & Gwizdka, 2007). In addition, received emails are created by other people, so they require more time to understand, evaluate, process and organize (Bellotti, Ducheneaut, Howard, Smith, & Grinter, 2005; Boardman & Sasse, 2004; Whittaker, 2005; Whittaker, Bellotti et al., 2006). Moreover, emails often require further actions including responding, keeping, organizing, deleting and deferring (Whittaker, 2011). Furthermore, iterations and delays between emails make managing emails difficult (Whittaker, Bellotti et al., 2006; Whittaker & Sinder, 1996). In addition,

people receive a high number of emails everyday so the problem of information overload is particularly severe in email management (Ducheneaut & Bellotti, 2001; Mackay, 1988; Whittaker, 2005).

2.3.2.3. Web Information

In the case of web information, some studies examined keeping and organizing behavior regarding bookmarks, and identified strategies used in keeping and organizing bookmarks (Abrams et al., 1998; Boardman & Sasse, 2004). In addition, there have been a number of studies that specifically focused on keeping activities (Abrams, Baeker, & Chignell, 1998; Bruce, Jones, & Dumais, 2004; Jones, Bruce, & Dumais, 2001; Jones, Dumais, & Bruce, 2002). In these studies, researchers found that people keep web information in a variety of ways. To be more specific, people kept web information by (1) sending email to themselves, (2) sending email to others, (3) printing out the web page, (4) saving the web page as a file, (5) pasting the URL for a web page into a document, (6) adding a hyperlink into personal web site, (7) bookmarking the web page, (8) writing down the notes on paper containing the URL of the web page, and (9) copying to a links toolbar to view the web address in the browser (Bruce et al., 2004; Jones et al., 2001; Jones et al., 2002). Thus, when keeping web information, people often converted it into other forms of information such as emails, papers, and electronic files. Among various ways of keeping web information, one of the most popular ways of keeping web information was creating bookmarks (Bruce et al., 2004; Jones et al., 2001; Jones et al., 2002).

2.3.2.4 Electronic Information

One of the main issues that has been addressed in management of electronic files has been whether people find it easier to search or browse to access personal information (Barreau, Capra, Dumais, Jones, & Perez-Quinones, 2008; Bergman, Beytg-Marón, Nachmias, Gradovitch, & Whittaker, 2008; Dumais et al., 2003; Jones, Phuwanartnurak, Gill, & Bruce, 2005; Teevan, Alvarado, Ackerman, & Karger, 2004; Whittaker, 2011). On the one hand, some researchers argue that people prefer searching to browsing when accessing personal information. Their main claim is that in the case of electronic personal information, organizing is becoming less important as search functions allow quick finding of information scattered across different locations in a personal device (Cutrell, Dumais, & Teevan, 2006; Russell & Lawrence, 2007). On the other hand, other researchers report that people prefer browsing to searching when accessing personal information in digital forms (Barreau & Nardi, 1995; Bergman et al., 2008; Bergman, Whittaker, Sanderson, Nachmias, & Ramamoorthy, 2010; Boardman & Sasse, 2004; Henderson, 2005; Jones et al., 2005; Teevan et al., 2004). Also, there are various functions of organizing personal information in addition to finding information, such as reminding tasks, or further understanding their information. A number of studies also found that people are still actively engaged in organizing their personal electronic files into organizational structures (Bergman et al., 2010; Boardman & Sasse, 2004; Gonçalves & Jorge, 2003; Henderson & Srinivasan, 2009). Thus, it is critical to understand organization of personal information in digital forms.

2.3.3 Organizing Personal Information

In the personal information management literature, there are a number of interesting studies that specifically examined the organization of personal information. In terms of personal information management, organizing activity is primarily an activity of grouping related information items together, and separating certain information items from other information items that are not related, based on certain criteria so that a person can easily understand and find information items. While a number of studies have investigated personal information organization, they rarely focused on the *process* of organizing personal information. However, they help explain the characteristics and various aspects of personal information organization. General findings from previous studies that examined organization of personal information can be grouped into six areas: types of personal information; organizational structures; changes in organizational structures; organizing strategies; organizing criteria; and factors that influence personal information organization.

2.3.3.1 Types of Personal Information

In studying personal information organization, some researchers identified different types of personal information, which is closely related to organization behavior. For instance, Cole (1982) identified three different types of personal information in the office: (1) action information, which is in current use or will be used in the near future; (2) personal work files, which are relevant to people's ongoing work; and (3) archive storage, which is not directly related to current work. Cole reported that action information is least formally organized, personal work files are more formally organized

in structures, and archive storage is the most formally organized (p. 61). Similarly, by integrating two separate studies on organizing personal information, Barreau and Nardi (1995) identified three types of personal information in the office: (1) ephemeral information, which is needed for a short period of time; (2) working information, which is frequently used information that is directly related to the participants' current work; and (3) archived information, which is archived for months or years and is indirectly related to the user's current work (p. 41-42). The researchers found that participants loosely organized ephemeral information, systematically organized working information and rarely organized archived information (p. 42). Thus, these studies show that personal information can be grouped into several types, and the way people organize their personal information is related to the different types of information items.

2.3.3.2 Organizational Structures

There have been several studies that investigated organization structures of personal information such as number of files, number of categories, number of subcategories, size of each category, and depth of organizational structures. For instance, Henderson and Srinivasan (2009), who examined the personal digital files of 73 knowledge workers, reported that on average people had 5,850 files, 628 folders, and 11.1 files per folder. In addition, on average, people had 3 subfolders per folder, and 6.8 depth of the folder structure (p. 397). The authors also found significant variation in the size and depth of folder structure (p. 402). These results are very similar to Gonçalves and Jorge's (2003) study which found almost the same mean number of files and individual variances. Nardi, Anderson and Erickson (1995) examined how people

organize and find electronic files on their computers by interviewing 15 Macintosh users (p. 1), and also reported various styles of organizing information in terms of labels, file names and folder structures (p. 1). Bergman et al. (2010) investigated folder structures of electronic files of 296 participants and found that they have an average of 22.46 files per folder, 10.64 subfolders, and 2.86 depth of the folder structure (p.2436). In a similar vein, Whittaker et al., (2010) also investigated 18 participants' organizational structures of photographs and reported that participants kept an average number of 4475 pictures, and usually had a single flat hierarchy and small number of subfolders (p. 11-12). Consistent with previous studies, this study also found a huge amount of variation in organizational structures among participants (p. 11). Boardman and Sasse (2004), who conducted cross-form studies, also reported that on average, people had 56.6 folders for electronic files, 32.3 folders for emails and 16.8 folders for bookmarks (p. 585). In addition, the average depth of the folder structure was 3.3 for electronic files, 1.7 for emails, and 1.3 for bookmarks (p. 585). Interestingly, 3% of electronic files, 41.6% of emails and 38.8% of bookmarks were not organized at all (p. 585). These results show that people organize electronic files more extensively than emails or bookmarks, and have more folders as well as deeper folder structures. Thus, people had a large number of personal information items stored in their personal devices, and categorized most of them into folder structures. In addition, size, number, and depths of the folder structure varied among people. Thus, these studies show that people are actively engaged in organizing personal information, usually by categorizing them into folders; they also show that in studying the process of personal information, individual differences need to be taken into account.

2.3.3.3 Changes in Organizational Structures

Gonçalves and Jorge (2003) examined the content and structure of personal information in digital form and observed changes made to organizational structures. In their studies, the authors found that compared to the total number of files they have, people created 0.9% of files per day, and 6% of files per month on average. In the case of modification, people modified 0.3% of files per day, and 3.7% of files per month to their already stored files (Figure 4).

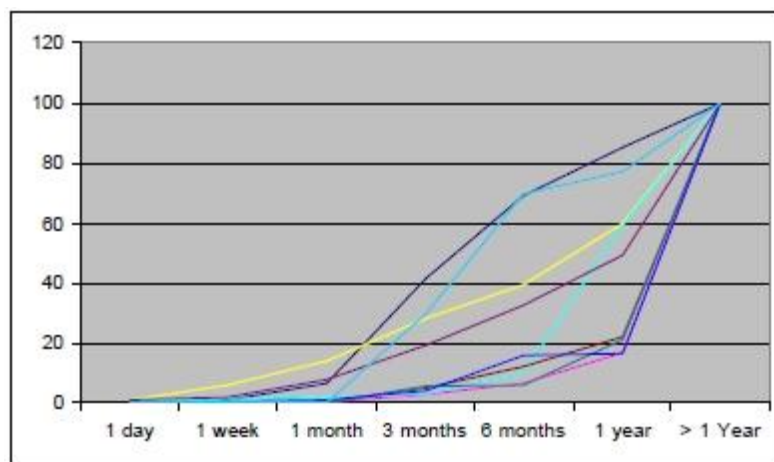


Figure 4. Percentage of Files by Modification Date (Gonçalves & Jorge, 2003)

Thus, although this study investigated changes occurring in files rather than folders, this study shows that personal information items are often newly created and modified. It also indicates that personal information organization changes over time. In a similar vein, Ravasio, Schär and Krueger (2004) examined personal information in electronic environments including electronic files, emails and bookmarks; they reported that participants thought of classification as an on-going process, and no structure was thought of as permanent (p. 164). This is also mentioned by Whittaker (2011) who stated that the process of organizing information is organic in that people often revisit and

restructure information to meet current needs (p. 9, p. 24). In addition, Ravasio, Schär and Krueger also found that participants create new folders when they feel it is important to keep an overview, or when there were three to seven files on the same subject (p. 164). While it was not a primary focus of the study, Barreau (1995) also briefly mentioned a participant who kept creating subfolders as the amount of information grew (p. 337). Thus, these studies indicate that personal information organization process is a dynamic on-going process.

2.3.3.4 Organizing Strategies

People often have ritualized ways of organizing information on their own (Nardi, Anderson & Erickson, 1995). Since Malone (1983) initially identified two different types of organizing strategies, a number of studies have identified various other types. Malone's study examined the ways office workers organize their desks and offices by observing and interviewing 10 office workers, and he found that there are (1) filers, who explicitly title and organize their information, and (2) pilers, who keep information without a systematic arrangement (p. 105). Based on this identification, Whittaker and Sidner (1996) identified three different types of organizing strategies by analyzing email organization. These strategies include (1) no filers, who do not organize their emails, (2) spring cleaners, who organize their emails periodically, and (3) frequent filers, who actively organize emails into folders on an ongoing basis (p. 281). Similarly, Bälter (1997) also identified and extended different types of organizing strategies by deleting 'no filers' and adding 'folderless cleaners' and 'folderless spring cleaners' to Whittaker and Sidner's three types of organizing strategies. Fisher, Brukh, Gleave, and Smith

(2006) also attempted to identify different organizing strategies for emails to compare their results to Whittaker and Sidner's study. However, the researchers found those three types of email organization behavior were difficult to identify (p. 312). Similarly, Dabbish, Kraut, Fussell, and Kiesler (2005), who attempted to categorize the email management strategies of 11 participants into three types of organizing strategies that were identified by Whittaker and Sidner, also failed to categorize them into Whittaker and Sidner's three types. Mackay (1988) identified two different types of organizing strategies for emails: (1) prioritizers, who organize emails as soon as they receive emails, and (2) archivers, who do not organize emails. Gwizdka (2004) also identified two types of organizing strategies: (1) cleaners and (2) keepers. While these studies identified organizing strategies for emails, Abrams et al. (1998) identified organizing strategies for bookmarks. They identified three different organizing strategies that are (1) no filers, who never organize bookmarks, (2) creation-time filers, who organize bookmarks when they are created, (3) end of session filers, who organize bookmarks at the end of the session, and (4) sporadic filers, who occasionally organize bookmarks (p. 45).

In terms of electronic files, Boardman and Sasse's (2004) cross-form study identified three different types of organizing strategies. These include (1) total filers, who organize most electronic files, (2) extensive filers, who organize many files, yet leave many files unorganized, and (3) occasional filers, who occasionally organize files and leave most files unorganized (p. 585). In terms of the timing of organization, the researchers stated that total filers tended to organize new items immediately (p. 585), while extensive filers organized items after finishing a relevant task (p. 586). However, the process itself was not specified. Boardman and Sasse also reported that the

participants tended to organize electronic files more extensively, with deeper folder hierarchies than emails or bookmarks (p. 585). These studies indicate that there are variations in time and frequency when people organize personal information. Thus, in studying the process of organizing personal information, it should be understood that some people might organize information right after they receive information while other people may organize information after a while. In addition, while some people decide to organize information, there will be other people who decide not to organize information items at all. Similarly, some people may organize information frequently, while others do not. Different organizing strategies identified in personal information management literature are summarized in Table 1 by format of information items.

Table 1
Organizing Strategies of Different Forms of Personal Information

Format of Information	Organizing Strategies
Paper-based Info.	<ul style="list-style-type: none"> • Filers, filers (Malone, 1983; Whittaker & Hirschberg, 2001)
Emails	<ul style="list-style-type: none"> • No filers/ spring cleaners/ frequent cleaners (Whittaker & Sidner, 1996) • Spring cleaners/ frequent cleaners/ folderless cleaners/ folderless spring cleaners (Bälter, 1997) • No filers/ spring cleaners/ frequent cleaners/ few folder filers (Fisher, Gleave & Smith, 2006) • Keepers/ cleaners (Gwizdka, 2004) • No filers/ partial filers/ extensive filers/ frequent filers (Boardman & Sasse, 2004) • Prioritizers/ archivers (Mackay, 1998)
Web Information	<ul style="list-style-type: none"> • No filers/ partial filers/ extensive filers (Boardman & Sasse, 2004) • No filers/ creation-time filers/ end of session filers/ sporadic filers (Abrams, Baeker & Chignell, 1998)
Electronic Files	<ul style="list-style-type: none"> • Occasional filers/ extensive filers/ total filers (Boardman & Sasse, 2004)

2.3.3.5 Organizing Criteria

There have been several studies which examined the criteria that are used in organizing personal information. For example, Case (1991), who examined how 20 historians organize paper-based information, found that type of information such as ‘books’ or ‘journals’ was one of the main criteria people use when organizing information items, followed by topic of information items (p. 662-663). Rodden (1999) investigated the organization of photographs, and reported that photographs are mostly organized by events and by time (p. 2). Rodden and Wood (2003) also observed that people mostly organize photographs by events and time. In the case of electronic files, Nardi, Anderson, and Erickson (1994) investigated 15 participants’ organization and retrieval of personal information in the office and found that people organize their files primarily by project or time (p. 3). Barreau (2008) also examined how 4 managers organize their electronic files and reported that people organize electronic files by task, topic, source, and form (p. 315-316). Thus, findings of these studies show that people use various criteria in grouping and separating information items including form, topic, source, task, and time.

2.3.3.6 Factors that Influence Organization

Since Kwasnik’s (1989) research, which investigated how 8 faculty members organize paper-based information in their offices, a number of studies have examined the factors that influence the organization of personal information. Kwasnik (1991) found 34 factors that impact organizing decisions, which included not only document factors such as author or topic of the document, but also various situational factors such as space or

time. In her study, form, use/purpose, topic, location, circumstance, and time were the primary factors that influenced organization decisions (p. 394). Based on Kwasnik's research, Barreau (1995) examined the factors that impact organization rules, and reported that various factors including topic, format, owner, currency, purpose, importance, intended use of information, and quantity and complexity of work affect organizing decisions (p. 333, p. 337). After 13 years, Barreau (2008) revisited the participants to see how advancing technology influenced participants' behaviors. Barreau reported that task, topic, provenance, and form influenced organizing behavior (p. 315-316), and found that behaviors had not changed much, which indicates that the impact of technology is less critical for how people organize personal information than their tasks are (p. 307). There have also been studies which specifically focused on the impact of goal on the construction of categories (Barsalou, 1983; Ratneshwar, Barsalou, Pechmann, & Moore, 2001). Barsalou stated that in our everyday life, people often construct 'ad hoc categories' which are highly specialized and unusual sets of categories that are constructed to achieve certain goals (p. 211). Ratneshwar et al. (2001) also found individuals' personal goals and situations influence perceived similarity of things (Ratneshwar et al., 2001). Thus, these studies show that there are a variety of factors that influence organizing decisions. More specifically, attributes associated with information items (e.g. topic, author (creator), or source of information item), attributes related to individuals (e.g. value, interest, importance), or attributes relevant to situations (e.g. the main use of information item and specific goals) all affect the organizing decision.

Thus, there are a number of studies that investigated personal information organization. As previously mentioned, these studies are helpful in understanding the

characteristics and various aspects of personal information organization. Especially, Kwasnik's study is closely related to this research study. However, most studies on personal information organization investigated the end results of organization, or focused on a specific aspect of the process rather than examining the whole process of organizing personal information. Therefore, although many studies revealed critical and interesting findings, in terms of understanding the process of organizing personal information, these studies provide an only partial view of personal organization.

To sum up, as shown above, a number of previous works in categorization, information behavior, and personal information management are relevant to this research study. These studies explored and found numerous critical and insightful findings in understanding organization. However, in the case of categorization theories, although categorization has been studied for many years, most of the early categorization theories examined categorization conceptually, or investigated classification of objects in the natural world rather than of information objects or personal information. In addition, how people organize their information was not a main focus of these studies. In the case of studies that examined the process of information behavior, the primary focus has been on the process of information seeking rather than organizing. In particular, there seem to be no studies which have investigated explicitly the process of organizing personal information. In the case of the personal information management literature, while there are a number of studies that investigated personal information organization, most studies investigated the end results of organizational structure or examined certain aspects of the process so that they provide a limited view of personal information organization. Most importantly, social influence on the organization process has been investigated rarely.

Thus, we still know little about how these organizational structures are constructed, what decisions are made, what is happening cognitively, and what factors impact people's grouping and separating information items during the process of organizing personal information. Thus, this dissertation suggests that a cognitive sociological model of the information organization process could explicitly address this issue.

CHAPTER 3 THEORETICAL FRAMEWORK

3.1 Categorization in Cognitive Sociology

In this study, cognitive sociology is used as the main theoretical framework. Cognitive sociology provides a unique and insightful framework for understanding how categories are constructed as well as cognitive aspects of the organization process, which is especially important in understanding the information organizing process. The viewpoint of cognitive sociology is that people think not only as individuals and as human beings, but also as social beings that are located in particular social environments (Zerubavel, 1997). Thus, in this study, the process of organizing personal information is viewed not only as an individual process and a universal process, but also as social process. The stance of cognitive sociology is presented in Figure 5.

Cognitive Individualism	Cognitive Sociology	Cognitive Universalism
Thinking as individuals	Thinking as members of thought communities	Thinking as human beings
Subjectivity	Intersubjectivity	Objectivity
Personal experience	Conventional cognitive traditions	Natural or logical inevitability
Personal cognitive idiosyncrasies	Cultural, historical, and sub cultural cognitive differences	Universal cognitive commonalities

Figure 5. The Stance of Cognitive Sociology (Zerubavel, 1997)

3.1.1 Lumping and Splitting

Cognitive sociology views classification not only as a personal act which people perform as individuals or as a universal act which people all perform as human beings,

but also as a social mental act that people perform as social beings (Zerubavel, 1997). In essence, society affects the way people classify the world (Zerubavel, 1991). Examples of the social nature of classification are certain distinctions that are made only in some societies but not in other societies, or modification of previous categorizations over time (Zerubavel, 1991). Zerubavel, a sociologist who studies classification from a cognitive perspective (DiMaggio, 1997, p. 276), stated that although many classification schemes seem either idiosyncratic or universal, often the way people categorize certain objects is quite similar to the way others around them categorize them (p. 53). In particular, Zerubavel uses the expressions ‘lumping’ and ‘splitting’ in explaining the process of category constructions. According to Zerubavel (1996), the world is continuous; however, we conceive it as discrete chunks such as ‘fiction’ or ‘nonfiction’ through the mental process of ‘lumping’ and ‘splitting’ (p. 421). In the process of ‘lumping’, people group things together by focusing on their similarities and overlooking differences among them. In the process of ‘splitting’, people separate things from one another by exaggerating their differences and ignoring similarities (Zerubavel, 1991; Zerubavel, 1996). Thus, while categorizing, people exaggerate intercluster mental distances and play down intracluster differences (Zerubavel, 1996); these mental activities are largely influenced by society, and especially individuals’ thought communities, which is explained in the following section.

3.1.2 Thought Community

A thought community (Denkgemeinschaft) is a community of people mutually exchanging ideas or maintaining intellectual interaction (Fleck, 1981). Any thought

community carries the historical development of particular fields of thought as well as the general stock of knowledge and culture (Fleck, 1981). However, it is important to note that thought community is not identical with official community (Fleck, 1981, p. 103). Examples of a thought community include a social class, political party, generation, nation, religious group, avocation, and occupation. Thus, people usually do not belong to one thought community but multiple thought communities at once (Fleck, 1981; Zerubavel, 1997). For instance, a person can be a doctor as well as a member of political party, religious group, nation, and so on. Thought communities shape people's perspectives and the ways people perceive the world around them that eventually make members of the communities perceive things similarly (Brekhus, 2010; Zerubavel, 1997). Thus, members of particular thought communities lump and split things in similar ways and often make distinctions between things which non-members would fail to notice (Zerubavel, 1996). This is because things are often considered as socially similar or different as people classify things as members of particular thought communities (Zerubavel, 1996). Interestingly, Fleck (1981) points out that an individual within the thought community hardly recognizes his/her thought style, despite its prevalence and its compulsive force upon his/her thinking. Thus, in studying the specific thought style of a thought community, a researcher needs to put a special effort in exploring the taken-for-granted thought style of the community. In this study, Fleck's 'thought community' is used as a core concept in exploring the process of personal information, especially in investigating the social aspects of the process.

3.1.3 Typifications

Berger and Luckmann (1966) argued that reality is socially constructed, and all human knowledge is developed, transmitted, and maintained in social situations; they did not focus directly on construction of categorization. However, their explanations about ‘typifications’ that occur during social interactions in everyday life provide some insights into understanding how people categorize things. Berger and Luckmann stated that during social interactions, people often apprehend others by typifying them (p. 29). For example, ‘my friend Henry’ can be typified as a member of category X, which is ‘an Englishman’. Then certain aspects of Henry’s behaviors such as his tastes in food or his manners can be interpreted as resulting from these typifications, i.e., as an Englishman (p. 30). Just as ‘my friend Henry’ can be categorized into ‘an Englishman’ through typifications, when we categorize an object, we typify it as ‘something’ so that it can be classified into a certain category. Cooper (2004) also stated that when we understand new information, we typify it according to previous experience and then put it into cognitive categories (p. 302). Thus, before categorization, typifications occur. Zerubavel (1991) stated that “the ability to ignore the uniqueness of items and regard them as typical members of categories is a prerequisite for classifying any group of phenomena” (p. 17). Here, we can also recognize that to typify an object as ‘something’, we need to focus on certain aspects of the object while disregarding other aspects of it. For instance, if a person categorized a certain book into a ‘not interesting books’ category, it means that he/she disregarded the book’s topic, genre, author, size, color, and publication year but typified it as a book that is not interesting based on his/her experience with other books that were not interesting to him/her, so that it can be categorized into a ‘not interesting

books' category. Thus, Berger and Luckmann's explanations about typifications help understanding how people construct categories, i.e., by typifications.

3.1.4 Anomaly and Ambiguity

Douglas' (1978) work on purity provides another meaningful insight into comprehending categorization. First, she claimed that 'dirt' is disorder, and thus, eliminating it is a positive effort to organize our environment (p. 2). Insightfully, she stated that the existence of 'dirt' means there is a system because 'dirt' cannot exist alone, i.e., 'dirt' can be only 'dirt' when there is a systematic ordering and classification of something (p. 35). In this sense, 'dirt' is either 'anomaly' or 'ambiguity' which violates or confuses existing classifications (p. 5). An 'anomaly' is "an element which does not fit a given set or series", which violates existing classification structures (p. 37); while 'ambiguity' is an element that can be interpreted in multiple ways, which confuses existing classification structures (p. 37). Thus, when people face an 'anomaly', people either ignore it by not perceiving it, or deliberately try to revise the existing scheme of classifications (p. 38). In a similar vein, when people face an 'ambiguous' object, people either reduce the ambiguity by interpreting it in a certain way, or physically eliminate it (p. 39). This observation is very useful in studying the process of organizing personal information for various reasons. First, it provides an explanation for why people initiate organization. When there are information items that are not organized, which are not parts of the existing categories, people may initiate organization in an effort to eliminate 'dirt'. Second, this work shows that both 'anomalous' and 'ambiguous' information items are difficult to categorize in contrast to the 'prototype' information items which were

previously discussed. Thus, we can hypothesize that those items that are difficult to categorize will be either anomalous or ambiguous items. Third, this work helps understand and predict how people will treat anomalous or ambiguous information items during the process of organizing personal information. For instance, we can hypothesize that when people encounter an information item that does not fit into one of the existing categories, i.e., an anomalous item, they may ignore it and not categorize it, or restructure categories by creating a new category or modifying existing categories. We can also hypothesize that when people confront an information item that can be categorized into multiple categories, i.e., an ambiguous item, they may categorize it into one of the existing categories or delete the item. Last, this work suggests the hypothesis that people are reluctant to revise their existing categories. Douglas (1978) stated that when people are categorizing, they simply accept those items that easily fit into the existing categories, and accept somewhat ambiguous items by treating them as if they are one of the members of the category. However, in the case of items that are so different from other items that they cannot fit into any categories, people often ignore or distort them so they do not have to modify their structure of assumptions which they have established and kept with confidence over time (p. 36). To sum up, this work provides a framework for understanding one of the motivations of initiating organization, why some items are difficult to categorize, how people deal with items that are difficult to categorize, and why people do not want to restructure their classifications.

As shown above, cognitive sociology's approach shows that in the process of categorization, people assess and adjust their mental distances between categories by ignoring or exaggerating similarities and differences. In addition, this approach highlights

the importance of recognizing the influence of society when studying the process of personal information organization. Thus, this perspective is used as the main framework of this research.

3.2 The Initial Model

The initial model for this research has been developed based on the literature and the researcher's analysis of the literature in categorization, information behavior, personal information management, and cognitive sociology. Thus, this is a newly developed model in this area of research. It is important to note that the specific elements in the model are examples of what can happen rather than comprehensive elements. This initial model was developed to be modified and extended while gathering and analyzing actual data. The initial model is presented in Figure 6.

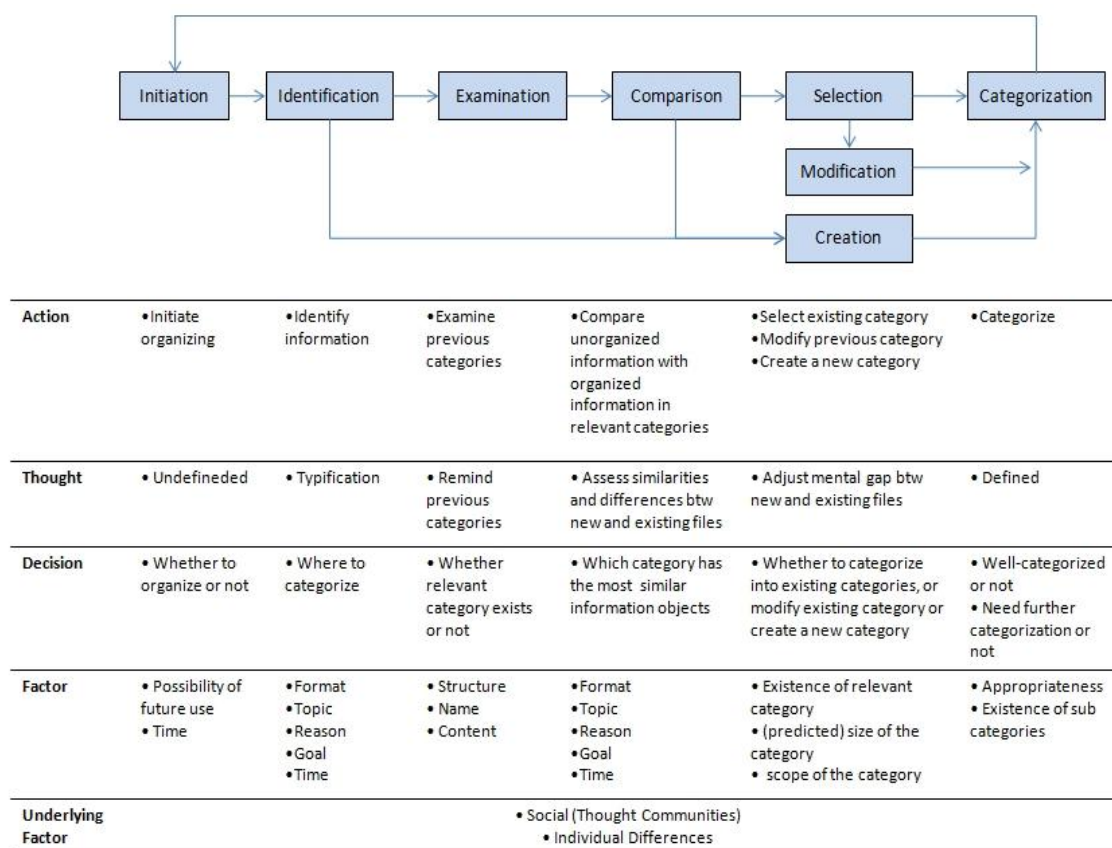


Figure 6. The Initial Model of the Personal Information Organizing Process (PIOP)

This model suggests that there are certain steps people take in organizing personal information, and there are common behaviors and thoughts people show in each stage of the process. Each stage also involves decision making. In addition, several factors influence the process. A detailed description of each stage follows.

3.2.1 Initiation

Depending on factors such as future use of information, a person may decide whether to save an information item or not. If an information item is something he/she is not going to use in the future, he/she will not save or organize this information item.

However, if there is a possibility of future use, he/she will save that information item, which initiates the organization process. At this point, how he/she will organize this information item, or how it will fit into his/her existing organization scheme is not yet specified.

3.2.2 Identification

To organize an information item into a category, he/she may focus on several aspects of an information item including its format, topic, time, or purpose. In this stage, he/she will typify this information item while focusing on certain aspect(s) of this information item. He/she may make typifications like “this is a PowerPoint slide” or “this is for my English Literature class”. In the process of such typifications, although each information item is unique, he/she transforms this information item into a somewhat familiar one (Zerubavel, 1997) so that he/she can identify and decide where to categorize that information item. Through this process of typification, his/her mental model is evolving, and he/she makes the information item fit into his/her world.

3.2.3 Examination

In this stage, a person will examine his/her existing organization schemes and the categories he/she created for his/her personal information. He/she may do this by physically viewing existing categories or simply by remembering the schemes or categories he/she has created. In this stage, he/she will examine the structures of the schemes, names of the categories, and information items that were already categorized in

those categories. However, this stage can be omitted if he/she simply creates a new folder despite the existence of categories that this information item may fit in.

3.2.4 Comparison

This stage is an extension of the previous stage, i.e., examination. When he/she physically checks what information items are in the existing categories or simply recalls existing categories, he/she will compare the new information item with existing information items by assessing similarities as well as differences between a new, unorganized information item and existing, organized information items in relevant categories so that he/she can decide which category has the most similar information items with the new information item that needs to be categorized. He/she will make this assessment based on various aspects of the information item such as format, topic, time, and purpose of the information item. However, this stage also can be omitted if he/she decides to create a new category for this information item without examining his/her existing categories, and if he/she does not notice that he/she already has a suitable category.

3.2.5 Selection/Modification/Creation

In this stage, he/she may select an existing category, or make modifications to the existing category if he/she thinks it will make this information item fit well in that category. In addition, there are several reasons why he/she may create a new category if he/she thinks a new category is needed. For instance, if there is no category that this information item can be categorized into, or an existing category has too many

information items, he/she will create a new category. Also, if he/she expects that there will be more information items to be organized into this category or if the existing category is too broad or too specific in scope, he/she will create a new category. In this stage, he/she will adjust his/her mental gaps between this information item and existing information items in existing categories. On the one hand, he/she will reduce the gap between this new information item and the information items in the existing category if he/she is going to select that category to organize an information item. On the other hand, he/she will inflate the gap between this information item and information items in other categories which he/she is not going to select in organizing the information item (Zerubavel, 1991).

3.2.6 Categorization

In this stage, he/she may place the information item into a category. However, if he/she thinks this information item is not well-organized, or that it needs to be categorized into other categories, he/she will initiate a new organization process. In addition, if he/she thinks it is necessary to categorize this information item into subordinate categories, he/she will also initiate another categorization process.

3.2.7 Underlying Factors

The theoretical framework for this study is cognitive sociology, and this model suggests that the social foundation of an individual who is organizing his/her personal information may affect the whole process of organizing personal information. To be more specific, the ways people identify and typify an information item that needs to be

organized, examine previous categories, compare similarities and differences between the to-be organized information item and already organized information items in relevant categories, and select the most relevant category or modify the previous category or create a new category, are all influenced by a social foundation and thought communities to which an individual belongs. However, it is important to note that individual differences are also one of the underlying factors. As people are not only social beings but also individual beings (Zerubavel 1997), it is assumed that people will have their own idiosyncratic ways of defining and organizing information items as well.

Consequently, this model suggests a framework for understanding the stages that constitute the process of organizing personal information, including behaviors and thoughts that are shown in each stage, decisions that are made in each stage, and factors that influence the process. In examining the process of organizing personal information, those six stages, common behaviors, thoughts, decisions, factors, and influence of the thought community will be the main focus of the examination.

3.3 Research Objectives

This study aims to explore the process of organizing personal information by identifying stages of the process, actions, thoughts, decisions, and factors that are involved in the process, and to then develop a model that explains that process holistically. The specific research questions are as follows:

RQ1. What are the stages of the personal information organization process?

RQ2. What actions do people take during the process of organizing personal information?

RQ3. What cognitive process do people go through during the process of organizing personal information?

RQ4. What decisions do people make during the process of organizing personal information?

RQ5. What factors influence the process of organizing information?

RQ6. How do individuals' thought communities influence the process of organizing personal information?

RQ7. Do participants always go through certain stages, actions, thoughts, decisions, and factors during the process?

CHAPTER 4 METHODOLOGY

4.1 Methods Used in Previous Work

To identify possible methods for collecting and analyzing data, and to investigate what types of data can be collected via certain methods, the various methods researchers used to attain similar types of data were investigated by reviewing some key previous works in personal information management and information behavior studies, with a special focus on their methodologies.

4.1.1 Grand (Guided) Tour Interview

A number of personal information management studies used grand tour interviews. In a grand tour interview, the researchers ask participants to give a tour of their information space such as their offices, desks, and computers while describing the organization of their personal information. More specifically, the researchers ask participants to explain what information is where and why it is placed in there. In this method, no structured questions are used during the interviews. In Malone's study (1983), the researcher used the grand tour interview with 10 people in their offices. Kwasnik (1989, 1991) also asked 8 professors to give grand tours of their offices. Similarly, Nardi, Anderson, and Erickson (1994) used grand tour interviews to study 15 employees, followed by a structured interview. Barreau (1995) also used grand tour interviews with 7 managers in their offices to examine factors that impact classification decisions in electronic environments (p. 331). This method seemed especially useful in collecting data about participants' personal information organizational structure, criteria

that are used in organizing personal information, and factors that impact on personal information organization.

4.1.2 Think Aloud

There were number of studies that used the think aloud method to collect data. Think aloud method asks participants to think aloud while solving a problem, and then analyzes the verbal protocol (Van Someren, Barnard & Sandberg, 1994). Think aloud method was mostly used together with the grand tour interviews. However, sometimes it was used without involving a grand tour interview. For instance, Rieh (2000) used think aloud method to collect data about people's decision making process in selecting information among multiple sources in the web, and the extent of information quality and cognitive authority people are concerned with when searching and evaluating information objects in the web (p. 74). Wang and Soergel (1998) also used think aloud method with 25 participants in developing a model of document selection. In their study, the researchers asked participants to think aloud while evaluating the information and making decisions to examine the cognitive processes underlying document selection (p. 119). This method seemed very useful in collecting data during the process, and especially when investigating participants' cognitive processes (Rieh, 2000; Van Someren, Barnard & Sandberg, 1994).

4.1.3 Semi-structured Interview

Several studies that explored the organizing behavior of personal information used semi-structured interviews in collecting data, so that the researchers asked

participants predetermined yet open ended questions (Krathwohl, 1997). For example, Cole (1997) interviewed 45 doctoral students and asked questions about recent information events to examine how knowledge structure is modified due to information (p. 55). Case (1986) also interviewed 60 professors to examine the number of books, journals, filing cabinet drawers, and stacks they used (p. 101-102). This method is often used in combination with other methods. For instance, after conducting grand tour interviews, Malone conducted semi-structured interviews to investigate how well participants felt their offices are organized, and what problems they faced when organizing personal information in their offices. Nardi, Anderson, and Erickson (1995) also conducted semi-structured interviews after grand tour interviews. In the case of Whittaker and Hirschberg's (2001) study, the researchers performed semi-structured interviews after conducting a survey.

4.1.4 Questionnaire

Some studies used quantitative research methods that involve surveying a large number of participants by using questionnaires. For example, after conducting an informal interview as a pilot study with 12 university students, Abrams, Baekcer, and Chignell (1988) surveyed 322 web users to examine the methods of organizing bookmarks, and the timing of organization (p. 45). Bälter (1998) also used a questionnaire to investigate problems with organizing emails, size of email flow, and perceived level of organization of files and papers. Similarly, Henderson (2009) used a survey to question 115 participants, along with interviews with 10 knowledge workers and snapshots of their file systems. In addition, Whittaker and Hirschberg (2001)

conducted an online survey with 50 workers, followed by semi-structured interviews with 14 workers to explore how people manage their personal paper-based information. As shown above, several researchers in personal information management used survey questions to collect data, and they often conducted surveys in combination with other data collection methods such as a pilot or follow up interviews. This shows the necessity of qualitative data in understanding personal information management behaviors, as these behaviors vary among people and require detailed explanations and interpretations.

4.1.5 Video Recording

Some studies video recorded participants' computer screens. For example, Bergman, Whittaker, Nachmias, and Ramamoorthy (2010) video recorded screens of 296 participants to investigate how they navigate and organize their personal electronic files. The researchers approached participants directly by knocking on the doors of students' residence halls; after printing out their recent document list, they asked participants to navigate each file. In terms of organization, by using video recordings, the researchers identified the depth of the folder structure, number of files per folder, and number of subfolders.

4.1.6 Diary Study

Some researchers used diary studies, which enables capturing data in a natural setting. Kuhlthau (1983) who developed the ISP model, collected data by using a diary study. The researcher asked 26 participants to record their actions, feelings, and thoughts about library search. In this study, the researcher asked participants to record in their

diaries what information resources they used as well as the procedures of finding those resources. In addition to the diary study, Kuhlthau also used a questionnaire which investigated participants' perceptions of six areas of library use, and conducted an interview with 6 participants to examine each stage of the process. Elswailer and Ruthven (2007) also conducted a diary study to identify what tasks lead to re-finding email and web information. In this study, 36 participants were asked to record details of their information re-finding tasks for three weeks (p. 25).

Thus, by reviewing some key literature, six possible data collection methods were identified: grand (guided) tour interview, think aloud, semi-structured interview, questionnaire, video recording, and diary study. After identifying possible methods to collect and analyze data, and the types of data that can be obtained from each method, the advantages and disadvantages of using each method are examined as shown in Table 2.

Table 2
Types of Data, and Advantages and Disadvantages of Using Each Method

	Types of Data	Advantages	Disadvantages
Grand (guided) tour interview	<ul style="list-style-type: none"> • Personal information organizational structure • Criteria used in organizing personal information • Factors that impact on personal information organization 	<ul style="list-style-type: none"> • Useful in understanding existing personal information organizational structure 	<ul style="list-style-type: none"> • Cannot collect data about the process of organization
Think aloud	<ul style="list-style-type: none"> • Cognitive aspects during the process • Decisions made during the process • Factors influencing the process 	<ul style="list-style-type: none"> • Useful in collecting data about decision making process • Useful in collecting data about cognitive process • Useful in collecting data about factors that influence the 	<ul style="list-style-type: none"> • Highly dependent on participants' verbalization skills • Cognitive processes are largely unconscious

		process	
Semi-structured interview	<ul style="list-style-type: none"> • Reasons for the action made • Reasons for the decision made • Factors that influence the decision 	<ul style="list-style-type: none"> • Useful in eliciting in-depth information about why certain decision was made during the process • Useful in eliciting in-depth information about why certain behavior was made during the process • Useful in eliciting in-depth information about factors that influence the decision 	<ul style="list-style-type: none"> • Participants need to recall their experiences • Highly dependent on participants' verbalization skills
Questionnaire	<ul style="list-style-type: none"> • Demographic information of participants • Criteria used in organizing information • Difficulties in organizing personal information 	<ul style="list-style-type: none"> • Useful in confirming pre-found data • Useful in collecting data that can be generalized 	<ul style="list-style-type: none"> • Cannot collect data about idiosyncratic behaviors • Cannot collect unexplored or unpredicted data • Cannot collect data that provide detailed interpretation • Participants need to recall their experiences
Video recording	<ul style="list-style-type: none"> • Actions made during the process • Current personal information organizational structure 	<ul style="list-style-type: none"> • Useful in collecting data about participants' behaviors • Useful in understanding existing personal information organizational structure 	<ul style="list-style-type: none"> • Privacy issues can occur
Diary	<ul style="list-style-type: none"> • Reasons for the decision made • Factors that influence the decision 	<ul style="list-style-type: none"> • Useful in collecting data in natural setting 	<ul style="list-style-type: none"> • Participants need to remember to record the diary • Highly dependent on participants' skills and dedication levels

4.2 Sources of Data

To collect data that address the research objectives, a short background questionnaire, a diary study, and two post hoc semi-structured interviews were selected as the most appropriate methods for this research study. Among various research methods, natural inquiry techniques such as observations, interviews, or diary studies provide powerful ways of understanding the contextual factors that influence individuals' personal information management styles in real-life settings (Naumer & Fisher, 2007). As this research aimed to discover the process of personal information organization and various decisions that are made during the process as well as the factors that impact on the process, this method seemed most appropriate.

4.2.1 Background Questionnaire

To collect some background information about participants, a short background questionnaire was chosen as one of the data collection methods for this study. Initially, in this questionnaire, some brief demographic information about participants such as gender, age group, ethnicity, and field of study were included. However, after the first pilot study, more questions were added to elicit information about participants' thought communities, which will be described in more detail in Section 4.4 and Section 4.6.

4.2.2 Diary Study

A diary study was chosen as it allowed the researcher to collect data which include various contexts of individuals from the perspective of participants. In addition, diary study allows studying behaviors over time rather than at one point in time. Also,

through diary study, participants can record details about their organizing activities which can be used later in conducting in-depth post hoc interviews. The diary also lets the researcher have some ideas about the organizing practice of each participant before conducting the interviews.

4.2.3 Interviews

Two separate post hoc semi-structured interviews were decided upon, because each interview can provide different information about participants' behaviors, including procedures and intentions. The first interview, which was conducted right after participants finished their recording diary, explicitly focused on the diary. The second interview, which was conducted 2-4 weeks after the first interview, focused on any changes that may have taken place with respect to the organization of those items discussed in the first interview. Thus, those two interviews performed different functions, and particularly the second interview was intended to examine the process of organizing personal information more holistically. As organization is an on-going and continuous process rather than a process that has a definite end (Ravasio et al., 2004; Whittaker, 2011), the second interview seemed necessary to collect data that provide a more comprehensive view of the process of organization of personal information. In fact, previous studies showed that people often revisit and modify their own classification schemes when the amount of information in a category grows or diminishes (Barreau, 1995; Gonçalves & Jorge, 2003; Ravasio, Schär, & Krueger, 2004) or when structure is out of date (Whittaker, 2011). Thus, it seemed necessary to include the second interview

to collect data that provide more comprehensive view of the personal information organization process.

4.3 Recruiting Participants

4.3.1 Recruiting Criteria

In this study, 18 participants were recruited from academics in a social science field in an institution of higher education in the United States. In particular, participants were selected from three sub-groups in different professional ages: undergraduate students; graduate students; and professors. In addition, the level of experience in each sub-group was considered when selecting participants. Gender and ethnicity were both balanced as much as possible. A detailed description about the criteria used in recruiting participants follows.

4.3.1.1 Thought Community

To explore how society, and especially thought communities, constructs people's cognition and influences the way people organize their personal information, a particular thought community has been chosen, i.e., academics. In this study, academics are defined as those who follow a university's academic calendar in their everyday life, which includes undergraduate students, graduate students, and faculty members. Academics are selected because academics deal with a number of information items in their everyday life by nature as most of their activities include interacting with information (Barreau & Nardi, 1995; Rieh, 2000). Academics not only create information items but also share, keep, and organize numerous information items. As the main purpose of the proposed

research is to examine the process of organizing personal information, academics were thought to be a good population for this study. In addition, because this is a case study which explores an area that has not been studied yet, it seemed appropriate to limit the scope of the study to a particular field of study, i.e., social science, before including and comparing different disciplines. For the same reason, the scope of the study is limited to institutions in the United States, before including and comparing institutions in different countries. Thus, participants were recruited from those who are in a social science field in an institution of higher education in the United States. In addition, since it is possible that ethnicity and gender differences may also affect the classification process, participants' gender and ethnicity were balanced as much as possible.

4.3.1.2 Professional Age

Professional age was used as one of the criteria in recruiting participants. In this study, professional age is defined as the number of years people have been engaged in higher education at any level. This criterion was considered because it is likely that the length of the time a person spent in certain thought community (i.e., academia for this study) will influence their organizing of information as a member of a thought community. For instance, undergraduates, graduates, and faculty can be different from one another, even though they are all academics. Thus, in this study, participants were selected from three sub-groups of academics: undergraduate students; graduate students; and professors.

In addition, it is important to note that there are different levels of experience in each sub-group of participants. In the case of undergraduate students, there are freshman,

sophomore, junior, and senior students who have spent different lengths of time in college. In the case of graduate students, there are master's students who focus more on professions, and doctoral students who focus more on research. In the case of professors, there are pre-tenured professors who have spent a relatively short time in higher education as a professor, as well as post-tenured professors who have spent a longer time. Thus, it seems necessary to consider these levels of experience in each sub-group of participants. In collecting data, to have participants who represent each sub-group, the researcher selected at least 2 participants in each representative class. Thus, for the 'undergraduate student' group, at least 2 participants were selected from (1) freshman or sophomore in college, and (2) junior or senior in college. For the 'graduate student' group, at least 2 participants were selected from (1) early research graduate students, and (2) late research graduate students. In the case of the 'graduate students' group, master's students were excluded from this study, as their primary focuses are on professions rather than research. For the 'professor' group, at least 2 participants were selected from (1) pre-tenured professors, and (2) post-tenured professors.

4.3.2 Recruiting Process

Participants were recruited by distributing a recruitment letter (Appendix 1) through email to each Social Science department so that those who were interested in this study could volunteer to participate. In addition, this research study was advertised in classes to undergraduate students, so that students could participate in this study as an extra credit opportunity. Also, the researcher's personal network was used to recruit participants.

4.4 Data Collection

The data collection process of this study is as follows. First, participants were introduced to the research study. Then, the informed consent form (Appendix 2) was given to the participants. They were asked to read the informed consent form, ask questions about the study if they had any, and sign the form if they agreed to participate in this study. In addition to the informed consent form, an audio/videotape addendum to the consent form was given to ask for participants' permission to allow the researcher to audiotape as part of the research study (Appendix 3).

Second, a background questionnaire was administered to collect background information about participants. The questionnaire can be found in Appendix 4.

Third, participants were asked to record diary entries on a given template over a week whenever they decided to (1) save a new information item in digital form onto any of their personal devices, or whenever they decided to (2) organize already saved information in their personal devices. In the diary template, the researcher provided a brief instruction including the definition of 'personal information', some detailed examples of instances which participants needed to record in the diary template, and the instruction about how to record a diary entry (Appendix 5). For each incident, participants were asked to record ID number, date, time, file name, format, source, reasons for saving, on which device they saved it, the path name if saved in an existing folder, the name of the folder if creating a new folder for the item, the names of the locations if moving items from one location to another location, and the names of the folders if changing the name of the folder. The template for the diary was created by the researcher in an Excel spreadsheet. (Appendix 6). The researcher also gave verbal

instructions about recording diary entries while looking at the diary template, and participants were asked to contact the researcher whenever they had questions about recording diary entries. While participants recorded their diary entries over a week, two reminder emails were sent to the participants. Also, participants were asked to send their diary templates after they were done with recording so that the researcher could take a look at it before the first interview.

Fourth, based on their diary entries, a first interview was conducted, which asked how and why participants saved and organized information (Appendix 7). This was a semi-structured interview, and participants were asked to recall and give explanations about how they saved and organized information and why they made such decisions while looking at the diary entries they wrote.

Fifth, about a month after the first interview, a second interview was conducted, which asked whether there had been any change to the files and folders that were discussed in the first interview (Appendix 8). A month seemed enough time for any changes to occur yet not so long that participants could not remember their organizing activities (Gonçalves & Jorge, 2003).

This data collection process is illustrated in Figure 7.



Figure 7. Data Collection Process

4.5 Data Analysis

The initial data analysis for this research study was conducted in the following way. First, all of the interviews were recorded and transcribed. Second, the researcher created tables for actions, thoughts, decisions, and factors for each stage of the process in a Word file. Third, from each transcript of the interview, the researcher identified and coded any actions, thoughts, decisions, and factors that belonged to certain stages of the process based on a set of categories (i.e., elements of the initial model) which were developed based on the literature and the researcher's analysis of the literature. While coding, the initial model was confirmed, extended, or modified. In this step, the researcher monitored if the data began to converge on a regular set of factors and their relationships to decisions including the influence of the thought community. Fourth, the researcher revised the initial model of the personal information organizing process.

4.6 Pilot Studies

Data collection and analysis methodology were investigated in two pilot studies.

4.6.1 First Pilot Study

4.6.1.1 Methods and Findings

The first pilot study was conducted to examine the effectiveness of the data collection instruments. In this pilot study, there were 3 participants, including 2 graduate students and 1 faculty member. From these 3 participants, three background questionnaires, three diaries, six interviews, and data on 58 information items were collected and analyzed.

In this pilot study, the diary and interviews were confirmed as usable and effective instruments in collecting data. However, the first pilot indicated that the background questionnaire as well as first interview questions needed some modification in order to elicit more data about the influence of the thought community on the organization process. Thus, necessary revisions were made to the background questionnaire and first interview questions.

4.6.1.2 Revisions: Background Questionnaire and First Interview Questions

In the case of the background questionnaire, originally the researcher only asked for some brief demographic information about participants (Appendix 4). However, the researcher added some questions which asked about participants' professional age, the primary roles they have in their daily life, and which role they regard as the most important source of identity for them. These questions were added to collect data to help investigate their thought communities and the influence of academic community on their lives as well as the organizing process. The revised version of the background questionnaire is attached in Appendix 9.

In the case of the first interview questions, the researcher added a question which asks about the primary devices the participants use in their daily lives, in order to have a better sense of their organizing behaviors. In addition, the researcher asked participants what they would do if they were not allowed to save this information item in the location in which they saved or organized each information items, so that they could consider other possible ways of organizing information items, which might eventually reveal some

information about why they made the current decision. The revised version of the first interview question is attached in Appendix 10.

In addition, the first pilot study revealed that it is important for the researcher to interview participants in a place where participants can easily access their personal information items. When participants could access their personal information items during the interview, they recalled more details about the context of saving and organizing information items they recorded in the diary. This also led them to explain about other information items they did not record in the diary, and allowed the researcher to have a better understanding about the organizing behavior of the participants.

4.6.2 Second Pilot Study

4.6.2.1 Methods and Findings

The second pilot study was conducted to test the adequacy of the revised instruments in collecting data for this study, with 2 further participants who were undergraduate students. From these 2 participants, two background questionnaires, two diaries, four interviews, and 29 information items were collected and analyzed.

The analysis of the data collected using the revised data collection instruments showed that they were effective in collecting more information about the influence of thought community on the personal information organization process. However, the second pilot also revealed that the data analysis method needed some modification in order to accurately understand the process of organizing each information item. In the case of the initial data analysis method, the researcher had fixed tables for each stage of the process. Then, the researcher tried to find evidence for actions, thoughts, decisions,

and factors that belong to a certain stage. Thus, in this data analysis method, the process of organizing ‘each information item’ was not considered. As a consequence, it was difficult to accurately identify stages of the process. Thus, necessary revisions were made to the data analysis method.

4.6.2.2 Revisions: Data Analysis Method

In the revised data analysis method, each information item was used as the basic unit of analysis so that the researcher could accurately identify stages and sequence of the process of organizing personal information. The detailed steps for the revised data analysis method are detailed next.

First, all of the interviews were recorded and transcribed. Second, the researcher added continuous line numbers in each transcript for coding. Third, the researcher identified each information item in the transcript because each information item was used as the unit of analysis in this study. Most of the time, it was easily identifiable from the transcripts by the file names, which represent those information items. Fourth, the researcher identified the location or category where the information item was saved or categorized. When it is categorized, it was mostly represented in the transcript as a folder name. Fifth, for each information item, any actions, thoughts, decisions, and factors that were shown during the process of organizing that information item were identified based on a set of initial coding categories. Then, identified actions, thoughts, decisions, and factors were typed in a coding scheme which was developed in Excel spreadsheets. This coding scheme allowed sorting data by each column, and facilitated finding data. A short description about each column of the coding scheme is as follows.

- (1) Participant number (P#)
- (2) Group of participant (Group)
- (3) Sub-group of participant (Sub-group)
- (4) Information ID (Info ID)
- (5) Name of the information item (File Name)
- (6) Thought community related with the information item (Thought Community)
- (7) Personal device where information item was saved (Device)
- (8) Format of the information item (Format)
- (9) Stage of the process (Stage)
- (10) Aspect of the process, which were actions, thoughts, decisions and factors
(Code level 1)
- (11) Detailed codes for each aspect (Code level 2)
- (12) Additional comments or details about each code if any (Code level 3)
- (13) Location of the quotation in a transcript (Line #)
- (14) Quotation from the transcript (Quote)

A part of the coding scheme is attached in Appendix 11.

Sixth, for each aspect of the process (i.e., actions, thoughts, decisions, and factors), the researcher identified stages of the process, and typed them in a coding scheme. During the fifth and sixth steps, the initial coding categories were confirmed, extended, or modified when necessary. When any changes were made to the initial coding categories, the codebook was revised accordingly and the transcripts were re-coded.

Seventh, after coding the entire set of transcripts, to examine relatively important actions, thoughts, decisions, and factors that are involved in the process of organizing personal information, the frequency of occurrences of each code were counted.

Eighth, revised codes were collapsed into broader codes by grouping similar codes so that very fine levels of distinction made with just few occurrences were grouped into more inclusive categories. However, although there were only few occurrences, if it was considered to be an important code to keep and make distinction from other codes, they were not grouped together. Again, in this process, the codebook was revised accordingly and the transcripts were re-coded. The final codebook is attached in Appendix 12. Ninth, the researcher created the personal information organizing process model for each information item. Tenth, the researcher integrated a number of models from the analyzed data and developed one model which best shows the ranges of behaviors that are shown during the process of organizing personal information. In the eleventh and final stage, to further analyze the social aspects of the process, the researcher analyzed each of the categories recorded in the diaries. This analysis process is explained in more detail in Section 5.2.6.

The reliability test was conducted to establish the reliability of the coding categories by using Stempel's Percentage Agreement Index (Stempel, 1955). In this process, two trained coders coded the same subset of data which were randomly selected from the transcripts of each sub-group of the participants. The Percentage Agreement Index for the subset of data was .85, which is the minimum (.85) commonly acceptable in social science research.

While coding data, the researcher also coded any part of the organization process of information items that were not recorded in the diary file, but were mentioned by participants. To be more specific, because participants were asked to record in a diary over a week whenever they decided to save or organize information items, the diary included information items that were saved over the past one week. However, when describing how they organized those information items, participants often mentioned already saved information items or existing folders when they were related to information items that were recorded in the diary. Participants also explained their organizing patterns or routines and other cases in which they organized information items in similar or different ways, and so on. Thus, in these cases, participants mentioned certain parts of the process without reference to a specific information item in the diary. Thus, in these cases, any identifiable parts of the process were coded. For example, in most cases, the researcher was able to elicit and code information such as (1) Participant number, (2) Group of participant, (3) Sub-group of participant, (9) Stage of the process, (10) Aspect of the process, (11) Detailed codes for each aspect, (13) Location of the quotation in a transcript (Line #), and (14) Quotation from the transcript (Quote). In addition, sometimes the researcher was able to code (6) Thought community related with the information item, (7) Personal device where information item was saved, or (8) Format of the information item when identifiable. However, it was impossible to code (4) Information ID. The researcher also coded any aspects of the process which had not taken place yet, but were mentioned by participants that they would happen in the future, or there was a possibility of happening.

In addition, the researcher added two different sheets in an Excel file so that the researcher could (1) code some interesting aspects of the organizing personal information in addition to the model, and (2) write notes while coding data.

CHAPTER 5 FINDINGS

5.1 Characteristics of Participants

Eighteen participants, who were academics in a social science field in an institution of higher education in the United States, participated in this study. In addition, participants were selected from three sub-groups with different professional ages: undergraduate students; graduate students; professors. In addition, as described in Section 4.3, two different levels of experiences in each sub-group were considered in selecting participants. Thus, the researcher selected at least 2 participants in six representative groups. The number of participants from each professional age group is presented in Table 3.

Table 3
Participants in Each Sub-group

Sub-group	Sub-sub group	Number of Participants
Undergraduates	Early	3
	Late	3
Graduates	Pre-qualifying Exam	3
	Post-qualifying Exam	3
Professors	Pre-tenured	2
	Post-tenure	4
Total		18

In analyzing and reporting the data, the researcher assigned an ID to each participant based on the sub-sub group of the participants. These IDs are presented in Table 4.

Table 4
Participants' ID

Sub-group	Sub-sub group	Participant ID
Undergraduates	Early	Early Undergraduate 1
		Early Undergraduate 2
		Early Undergraduate 3
	Late	Late Undergraduate 1
		Late Undergraduate 2
		Late Undergraduate 3
Graduates	Pre-qualifying Exam	Pre-qual Graduate 1
		Pre-qual Graduate 2
		Pre-qual Graduate 3
	Post-qualifying Exam	Post-qual Graduate 1
		Post-qual Graduate 2
		Post-qual Graduate 3
Professors	Pre-tenured	Pre-tenured Professor 1
		Pre-tenured Professor 2
	Post-tenure	Tenured Professor 1
		Tenured Professor 2
		Tenured Professor 3
		Tenured Professor 4

In addition, to protect the identity of participants, any information that might reveal the identity of the participants such as name of the university, department, conference, journal, or course were referred as 'University 1' or 'Conference 1', etc. When it was necessary to make differentiation between them, different numbers were assigned such as 'University 1' and 'University 2'. In addition, any person's name that was mentioned by the participants was also referred to by the relationship with that person such as 'student', 'professor', or 'colleague' rather than their unique names.

5.1.1 Demographic Information

The researcher tried to balance the number of participants from different genders as much as possible. However, most of the participants who volunteered to participate in

this study were women. As shown in Table 5, among 18 participants, 13 participants (72.2%) were female and 5 participants were male (27.8%).

Table 5
Participants' Gender in Each Sub-group

Sub-group	Sub-sub group	Gender	
		Male	Female
Undergraduates	Early	1	2
	Late	1	2
Graduates	Pre-qualifying Exam	1	2
	Post-qualifying Exam	1	2
Professors	Pre-tenured	1	1
	Post-tenure	0	4
Total (%)		5 (27.8)	13 (72.2)

The age groups of participants varied. Among 18 participants, there was one participant who was a teen (5.6%), 4 participants who were in their 20s (22.2%), 7 participants who were in their 30s (38.9%), 1 participant who was in their 40s (5.6%), 3 participants who were in their 50s (16.7%), and 2 participants who were older than 60 (11.1%). The number of participants in each age group is presented in Table 6.

Table 6
Participants' Age Group

Age Group	Number of Participants
10s	1
20s	4
30s	7
40s	1
50s	3
60+	4
Total	18

In the case of ethnicity of participants, among 18 participants, 1 participant was African American (5.6%), 6 participants were Asian (33.3%), 1 participant was Hispanic (5.6%), and 10 participants were Caucasian (55.6%). Table 7 shows the number of participants in each ethnicity.

Table 7
Participants' Ethnicity

Ethnicity	Number of Participants
African American	1
Asian	6
Hispanic	1
Caucasian	1
Total	18

As previously mentioned, participants were recruited from social science fields in an institution of higher education in the United States. Thus, participants were from social science departments. Of the total participants, 4 participants were from Communication (22.2%), one participant was from East Asian Studies (5.6%), one participant was from Economics (5.6%), one participant was from Education (5.6%), one participant was from Human Computer Interaction (5.6%), 2 participants were from Human Resource Management (11.1%), one participant was from Information Technology and Informatics (5.6%), 5 participants were from Library and Information Science (27.8%), one participant was from Planning and Public Policy (5.6%), and one participant was from a Sociology department (5.6%). Table 8 shows the number of participants in each social science department.

Table 8
Participants' Department

Department	Number of Participants
Communication	4
East Asian Studies	1
Economics	1
Education	1
Human Computer Interaction	1
Human Resource Management	2
Information Technology and Informatics	1
Library and Information Science	5
Planning and Public Policy	1
Sociology	1
Total	18

5.1.2 Thought Communities

As described in Section 5.1.2, a thought community is a group of individuals who share ideas, concepts, and theories (Erickson, 2005). Each of the thought communities has its own thought style, as well as its own way of making sense of and understanding the world. Thought communities vary in size, and people often belong to multiple thought communities (Fleck, 1981; Zerubavel, 1997). In this study, because participants were recruited from academics, ‘academia’ was one of the thought communities of all 18 participants. However, because participants were selected from different professional ages, the year spent at universities ranged from 1.5 to 46 years. The average years participants spent at universities were 2.4 years for undergraduate participants, 10.3 years for graduate participants, and 29.3 years for professor participants. The average year that all 18 participants spent at universities was 14 years. Table 9 shows the professional age of each participant.

Table 9
Participants’ Professional Age

Sub-group	Sub-sub group	Participant ID	Professional Age
Undergraduates	Early	Early Undergraduate 1	1.5
		Early Undergraduate 2	1.5
		Early Undergraduate 3	1.5
	Late	Late Undergraduate 1	2.5
		Late Undergraduate 2	3
		Late Undergraduate 3	4.5
Graduates	Pre-qualifying Exam	Pre-qual Graduate 1	10
		Pre-qual Graduate 2	7
		Pre-qual Graduate 3	9
	Post-qualifying Exam	Post-qual Graduate 1	13.5
		Post-qual Graduate 2	11
		Post-qual Graduate 3	11
Professors	Pre-tenured	Pre-tenured Professor 1	16
		Pre-tenured Professor 2	20
	Post-tenure	Tenured Professor 1	46
		Tenured Professor 2	32
		Tenured Professor 3	27

Although participants' professional age varied, almost all of the participants regarded the academic thought community as a very important source of identity for them. When asked whether being an academic is an important source of identity for them, participants said:

“School comes before everything for me.” (Early Undergraduate 1)

“When I think of myself and people ask me what I am, I’m like I’m a grad student. That’s really my life.” (Post-qual Graduate 2)

“I always identify myself as an academic.” (Pre-tenured Professor 2)

“It has been an incredible source of identity for me.” (Tenured Professor 1)

When participants were asked to explain why being an academic is an important source of identity for them, 10 participants (55.6%) across three sub-groups (i.e., undergraduates, graduates, and professors) mentioned about spending much time as an academic. For instance, participants said:

“Most of the year, I’m in school. And I’m like already studying; homework takes up most of my time.” (Early Undergraduate 2)

“I’m a full time student, so like, my mentality is, like, I guess, everything I do is pretty much like, surrounded around school, like homework, exam, like, um, yeah, I think in general, everything I do is like, really just school.” (Late Undergraduate 2)

“It’s the one that’s the most time consuming at this moment.” (Post-qual Graduate 3)

“It has a daily influence on my life.” (Tenured Professor 2)

Some undergraduate and graduate participants responded that they regard being an academic as a very important source of identify for them because it is connected to the future. Participants said:

“I feel like, school is for now, but it’s also for the future. And like what I’m going to do in the future, um, you know, with my job and everything, so, yeah. It’s more long term.” (Early Undergraduate 1)

“It’s a part of succeeding, and it’s a part of like my future and my life.” (Late Undergraduate 1)

“I think my current position as a student is really important to decide my career in the future.” (Post-qual Graduate 1)

A unique answer from undergraduate participants about why the academic community was important in their lives was parents’ expectations. For instance, participants said:

“School comes before everything for me. And like, my parents also like try to like enforce that.” (Early Undergraduate 1)

“I think school has always been important to me. Um, because like, how I was raised, how Asian parents raise, you know, school, like, very important?” (Late Undergraduate 2)

On the other hand, several graduate and especially professor participants mentioned the value and meaning of the academy in society, their pride as an academic, and why they like to be an academic, while no undergraduate participants mentioned this. For example, graduate students and professors said:

“I am proud of being associated with the institution and what I’m doing. I believe in higher education and I think it’s an important thing to be doing” (Pre-qual Graduate 2)

“I’ve been studying it for ten years and it means a lot to me.” (Pre-qual Graduate 3)

“I feel that by being an academic I can help others tremendously. For example, I can give graduate students, especially those from other countries, guidance in how to be successful academics and researchers in this competitive field. I can also argue for and create new courses that serve students better, I can give guidance, encouragement and advice to all levels of students and I can work on research projects that will generate results that will be extremely helpful to people.”

(Tenured Professor 1)

“Being a professor, we already talked, you know, it’s a part of, it reflects my values and who I am, and my belief in knowledge and my belief in learning, and passing it on to the next generations.” (Tenured Professor 3)

Thus, for undergraduate students, the importance of academic community was more influenced by external factors such as preparing for the future, or meeting the expectation of their parents rather than the academic community itself. However, for professors, it was more about the value and the meaning of the academic community for both society and themselves. The graduate student participants were in the middle of the road.

As described in Section 3.1, people usually belong to multiple thought communities (Fleck, 1981; Zerubavel, 1997). Thus, although academia is one of the participants’ thought communities, participants also belonged to other thought

communities that were non-academic as well. Participants' non-academic thought communities included family, friends, occupation, volunteer organization, political party, couple, nationality, and religious group.

To examine participants' primary thought community among multiple thought communities where participants belong, the researcher asked which role they regard as the most important source of identity for them. To this question, 14 participants (77.8%) out of 18 participants answered that academia is their primary thought community.

Among these 14 participants who regarded academia as their primary thought community, 3 participants said family or religious group is an equally important thought community to them. Among 4 participants (22.2%) who said being an academic is not the most important source of identity for them, 2 participants were part-time students (Early Undergraduate 3, Late Undergraduate 3), and they regarded their occupation as the most important source of identity for them. In addition, one of the graduate students (Pre-qual Graduate 2) answered that he/she regards being a member of a family as the most important source of identity for him/her. Another graduate student (Pre-qual Graduate 3) responded that his/her religious group is the most important source of identity for him/her.

5.2 The Process of Organizing Personal Information

From 18 participants of this study, 18 diaries, 36 interviews (18 participants x 2 interviews), and 235 diary entries were collected. The average number of diary entries each participant kept was 13.1, and it ranged from 5 to 37. The number of diary entries from each participant is shown in Table 10.

Table 10
Number of Diary Entries

Sub-group	Sub-sub group	Participant ID	Number of Diary Entries
Undergraduates	Early	Early Undergraduate 1	10
		Early Undergraduate 2	9
		Early Undergraduate 3	16
	Late	Late Undergraduate 1	12
		Late Undergraduate 2	10
		Late Undergraduate 3	12
Graduates	Pre-qualifying Exam	Pre-qual Graduate 1	16
		Pre-qual Graduate 2	8
		Pre-qual Graduate 3	11
	Post-qualifying Exam	Post-qual Graduate 1	11
		Post-qual Graduate 2	13
		Post-qual Graduate 3	17
Professors	Pre-tenured	Pre-tenured Professor 1	23
		Pre-tenured Professor 2	8
	Post-tenure	Tenured Professor 1	37
		Tenured Professor 2	5
		Tenured Professor 3	12
		Tenured Professor 4	5

5.2.1 Stages

RQ1. What are the stages of the personal information organization process?

As introduced in Section 3.2, the initial model, which was newly developed in this study based on the researcher's analysis of the literature, had six stages that constitute the process of organizing personal information. These six stages and brief descriptions about each stage are as follows:

- (1) Initiation: The first stage of the process, in which a participant decides to save an information item.
- (2) Identification: The second stage of the process, in which a participant identifies an information item in his/her own way so that he/she can organize an information item.

- (3) Examination: The third stage of the process, in which a participant scans his/her existing folders to find a relevant folder for the information item to be organized.
- (4) Comparison: The fourth stage of the process, in which a participant compares an unorganized information item with organized information items in the relevant folder.
- (5) Selection/Modification/Creation: The fifth stage of the process, in which a participant selects one of the existing folders, modifies one of the existing folders, or creates a new folder to organize an information item.
- (6) Categorization: The last stage of the process, in which a participant organizes an information item into a folder.

These stages are sequential, and because information organization is a dynamic process that may change over time (Gonçalves & Jorge, 2003; Ravasio et al., 2004; Whittaker, 2011), it was assumed that a participant might restart the organization process from the initiation stage to re-categorize an information item when necessary. These stages as well as the sequence of the process in the initial model are presented in Figure 8.

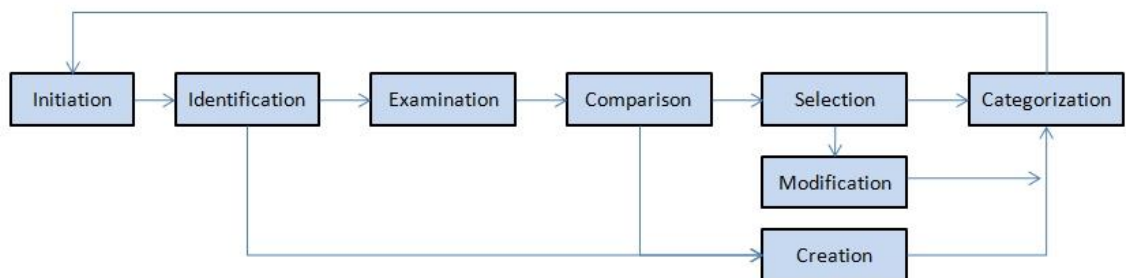


Figure 8. Stages of the Personal Information Organization Process in the Initial Model

Based on these stages of the process in the initial model, the researcher identified different stages of the personal information organization process while analyzing the empirical data. When analyzing data, any actions, thoughts, decisions, and factors that were involved in the process of organizing personal information were coded first, and then for each aspect of the process (i.e., actions, thoughts, decisions and factors), the stages were identified. In this process, some of the stages in the initial model were confirmed while other stages were modified. In addition, a new stage was found. To be more specific, (1) Initiation, (2) Identification, (5) Selection/Modification/Creation, and (6) Categorization stages were confirmed as parts of the process while analyzing the empirical data.

However, (3) Examination and (4) Comparison stages in the initial model were merged together into the (4) Examination/Comparison stage because often it was difficult to separate these two stages. In particular, examining existing categories and comparing an unorganized information item with organized information items in relevant categories almost always happened together. For example, participants said:

“It’s just that I have three sections, basically, of the same course. They’re all Course 1: one is grad, one is undergrad, and one is at University 1, so I’m saving them all to ‘Fall Courses’ because they’re three different courses, but they’re really the same course.” (Tenured Professor 3)

“So I have ‘Reviews’, and under ‘Reviews’ I have all the years I’ve been reviewing.” (Tenured Professor 1)

As shown in these examples, when participants reviewed relevant categories from the existing categories, they reviewed already organized information items in those categories at the same time. Here, their judgments about the ‘relevant folder’ were possible because they compared unorganized information item with organized items in those folders. Thus, it was almost impossible to differentiate which part of the description indicates reviewing his/her existing categories, and which part shows comparing an unorganized information item with organized information items in the category.

In addition, while analyzing data, a new stage, (3) Temporary Categorization stage was found. In this stage, participants delayed decisions and just saved an information item in a certain place for a short period of time. For example, participants said:

“I just temporarily put it on the ‘Desktop’.” (Pre-tenured Professor 2)

“I keep it in my ‘Downloads’ folder for the time being.” (Pre-tenured Professor 1)

Because this stage was an unexpected stage, at first this stage was coded as (6) Categorization stage. However, in these cases, participants did not really examine their existing categories before deciding to save information items into this temporary place. Thus, it was different from the (6) Categorization stage. In addition, a temporary location was different from a miscellaneous category in which participants saved information items that did not belong to any category. For example, one of the participants who had a ‘Miscellaneous’ folder described it as:

“I have one called ‘Miscellaneous’, current miscellaneous. It’s kind of anything that doesn’t fit.” (Tenured Professor 3)

Instead, participants often mentioned that they will re-organize those information items into their existing folder structures or delete the file at a later point. For example, participants said:

“I dump everything on ‘Desktop’, and then I file later.” (Tenured Professor 3)

“Usually I just save it to the ‘Desktop’ and then once I remember, I’ll put it in the designated folder.” (Late Undergraduate 3)

“I always move (files) from ‘Download’ to folders.” (Pre-tenured Professor 2)

In addition, when they said that they would re-organize those information items, they often reviewed some relevant categories where those information items could be organized. Participants described that:

“I saved it to my ‘Desktop’, which I will move into my ‘Spring Semester’ folder.”
(Late Undergraduate 1)

“And everything should go under ‘Fall Courses’, I just haven’t moved them yet.”
(Tenured Professor 3)

“I’ll put it in a folder for the ‘Summer session 2012’.” (Late Undergraduate 3)

In addition, several participants mentioned that there is a temporary place where they save information items which is separate from their folder structures. For instance, participants said:

“It’s like short-term memory and long-term memory. Short-term memory I can manage everything that is in the same folder because it’s easy.” (Post-qual Graduate 3)

“It’s like a temporary system that I then integrate with the permanent system.”
(Tenured Professor 3)

“‘Desktop’ is the really short term use, rather than a long term use, but I use my web site as my long term use, the files that I want to save for a long period of time.” (Post-qual Graduate 1)

Thus, it was different from the (6) Categorization stage, and it seemed necessary to revise the initial stages of the organizing process, and include the new stage ‘Temporary Categorization’ into the process. This new stage, ‘Temporary Categorization’, was included between (2) Identification stage and (4) Examination/Comparison stage because participants saved information items in temporary places after initiating the organizing process. In addition, participants identified information items before they saved them into a temporary place, so that not all information items were saved into a temporary place. Also, participants saved information items before examining their previous categories, but then often examined them later on. Thus, after (2) Identification stage and before (4) Examination/Comparison stage seemed the most appropriate place for the (3) Temporary Categorization stage. After revising stages of the process, all of the interview transcripts were re-coded. The revised coding scheme made it much easier to code the transcripts, which indicated that it reflects the process better than the initial coding scheme of the stages.

As a result, it was found that the process of organizing personal information has following six stages:

- (1) Initiation: The first stage of the process, in which a participant decides to save an information item.

- (2) Identification: The second stage of the process, in which a participant identifies an information item in his/her own way so that he/she can organize an information item.
- (3) Temporary categorization: The third stage of the process, in which a participant delays any organizational decision and saves an information item into a certain location temporarily.
- (4) Examination/Comparison: The fourth stage of the process, in which a participant scans his/her existing folders to find a relevant folder for the information item to be organized.
- (5) Selection/Modification/Creation: The fifth stage of the process, in which a participant selects one of the existing folders, modifies one of the existing folders, or creates a new folder to organize an information item.
- (6) Categorization: The sixth stage of the process, in which a participant organizes an information item into a folder.

The stages of the process in the initial model and the stages of the process that were found after analyzing empirical data are displayed in Table 11.

Table 11
Changes to the Stages of the Process

Stage	The Initial Model	The Final Model
Stage 1	Initiation	Initiation
Stage 2	Identification	Identification
Stage 3	Examination	Temporary Categorization
Stage 4	Comparison	Examination/Comparison
Stage 5	Selection/Modification/Creation	Selection/Modification/Creation
Stage 6	Categorization	Categorization

However, during the process of organizing personal information, not every information item went through all six stages. In other words, not all stages always appeared as a part of the process. Some information items went through only part of the process, such as (1) Initiation, (2) Identification, and (3) Temporary Categorization. For instance, a participant said:

“That was for a project I had to do for my strategic presentation class... As you can see, I just left it in ‘Downloads’.” (Late Undergraduate 1)

Also, as frequently as information items were just saved into a temporary category, there were many cases when information items were directly saved into the organizational structure without going through the (3) Temporary Categorization stage. For instance, a participant said:

“I downloaded the paper... [I saved it] in the laptop, and existing folder called ‘Class Name 1’ because during my semester, I created a folder called ‘Class Name 1’. I tried to save different files such as my reference paper or different, possibly guideline, or other Word documents. So I saved that final paper into existing folder to organize my whole class wise folder.” (Pre-qual Graduate 1)

Sometimes, because moving from one stage to another stage happens very quickly, continuously, and often unconsciously, certain stages were not discernible or not

mentioned by the participants. In particular, selecting an existing category in the (4) Selection/Modification/Creation stage and placing a file into a category in the (5) Categorization stage were often hard to differentiate in the verbal description of the participants. When participants said they put certain information item into a certain folder, it included both selecting one of the folders as well as placing an information item into a folder. For instance, participants said:

“So I put it into ‘Department name’ (folder) because I still have to read it.”

(Tenured Professor 4)

“So I just put it all in the same folder.” (Early Undergraduate 1)

However, in most cases, the researcher was able to identify different stages of the process.

As shown in the model, these processes were sequential. In addition, as previously mentioned, because information organization is a dynamic process that may change over time, in the original model, it was assumed that after a participant categorized an information item into a category (i.e., (6) Categorization stage), a participant might go back to the (1) Initiation stage to re-categorize an information item when necessary. Indeed, the analysis of the data showed that information items were often re-categorized after categorization. Sometimes participants actually re-categorized an information item, and sometimes they mentioned that they would re-categorize it in the future. For instance, participants said:

“I put the ones that have to do with one class into one folder. So I did that for a

few of my classes.” (Early Undergraduate 1)

“I actually at the end of the semester will just combine all my folders because I don’t need them for classes anymore.” (Late Undergraduate 1)

However, in the case of the sequence of the re-categorization process, the original model has been revised. In the original model, it was assumed that participants will restart the organization process from the beginning when they re-categorize an information item. However, the analysis of the data showed that when participants re-categorize an information item that was saved into a temporary location in the (3) Temporary Categorization stage, they were re-categorized into one of the folder structures so that it goes to the (4) Examination/Comparison stage. However, when re-categorizing an information item that was placed into a folder in the (6) Categorization stage, participants do not go back to the very first stage of the process (i.e., (1) Initiation). Instead, because participants already identified an information item, participants often examined previous categories, which is (4) Examination/Comparison stage, and then selected another existing category, or modified their previous categories, or created a new category, which is (5) Selection/Modification/Creation stage. Thus, the arrows of the model have been modified to reflect these findings. Examples of this re-categorization process are as follows:

“I have like a past folder yeah of other St. Patty’s Days, and I’ll just combine this [‘St. Patty’s Day 2012’ folder] in with that when I’m done with this.” (Late Undergraduate 1)

“I actually created a new folder called ‘Travel picture’ folder and then I put it [‘Summer Travel Picture 2012’ folder] there. I put the folder into the bigger category. Because I realized that there are many different traveling pictures

folders lying around in ‘My Pictures’ folder so I just wanted to collect them together in one folder.” (Pre-qual Graduate 3)

However, not all information items were re-categorized. Participants sometimes kept an information item in the category, or deleted an information item. For instance, participants said:

“I haven’t changed anything for that folder” (Pre-qual Graduate 1)

“No, it’s not gonna get moved.” (Post-qual Graduate 2)

“I actually deleted the file.” (Late Undergraduate 3)

These re-categorization decisions and factors that have an impact on the decisions are described in more detail in Section 5.2.4. and Section 5.2.5.

The stages and the sequence of the final model in this study are presented in Figure 9.

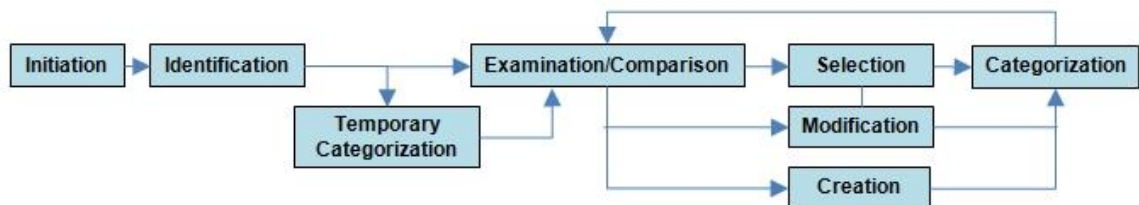


Figure 9. Stages of the Personal Information Organization Process in the Final Model

5.2.2 Behavioral Aspect

RQ2. What actions do people take during the process of organizing personal information?

One of the aspects examined when exploring the process of organizing personal information was the behavioral aspect, which refers to any actions participants show

during the process of organizing their personal information. While analyzing data, any actions that were involved in organizing personal information were identified and coded. Codes for actions that were developed while analyzing data as well as occurrences of each code are organized by each stage of the process and presented in Table 12. When counting occurrences, if the same code occurred multiple times in describing the process of organizing the same information item, they were counted as one occurrence. However, if different codes occurred in describing the process of organizing the same information item, each was counted separately. There were some cases when participants mentioned actions that were associated with information items that were not recorded in the diary, or future actions they would take, or possible actions they might take. These cases were all coded and sometimes presented as examples; however, they were not counted in the occurrences. Thus, only those occurrences that actually happened while organizing information items that were recorded in the diary are counted.

Table 12
Codes for Behavioral Aspect of the Process

Stage	Code	Occurrence
(1) Initiation	Receive file	21
	Create file	25
	Save file	120
	Obtain file in another way	5
(3) Temporary Categorization	Save file in a temporary location	42
(5) Selection/Modification/Creation	Select a folder	109
	Modify previous folder	15
	Create a new folder	16
(6) Categorization	Place file into a folder	115

In the (1) Initiation stage, participants started an organizing process when they had an information item. For example, when participants received a file from someone, saved a file, or created a new file, the information organizing process began.

Receive file. In the case of receiving files from someone, participants said:

“[I] finally got all files from that people by email.” (Pre-qual Graduate 1)

“So the files came together in the same email.” (Pre-qual Graduate 3)

Because the scope of information items was limited to the electronic format, mostly participants received an information item via email.

Save file. However, as shown in Table 12, when describing how they initiated the organizing process by having an information item, most often participants mentioned ‘saving files’ into their personal devices. For example, participants said:

“I just saved it to my ‘Desktop’.” (Early Undergraduate 2)

“I saved in the ‘Midterm’ folder.” (Post-qual Graduate 2)

When describing saving information files, among 120 occurrences, 87 occurrences used the expression ‘saving file’ while 33 occurrences used the term ‘downloading file’. For example, participants said:

“I just directly downloaded it from the web site.” (Pre-qual Graduate 1)

“I downloaded this.” (Tenured Professor 2)

Often participants downloaded information items from the web, databases, or emails.

Saving files not only included downloading files but also included taking a screen shot of a web site, uploading a file to a web site, or bookmarking a web site. For instance, participants said:

“I took the screen shot of it.” (Early Undergraduate 2)

“It automatically uploads anything new.” (Early Undergraduate 3)

“Any of the articles that I found really interesting I would basically open, bookmark and then I save them.” (Early Undergraduate 3)

Create file. Sometimes, participants initiated an organizing process by creating a new file. For example, participants mentioned:

“That’s another set of minutes that I created.” (Pre-qual Graduate 2)

“So I combine my lecture notes and my textbook notes into like one.” (Late Undergraduate 2)

Obtain file in another way. In addition, there were six occurrences in which participants used other ways of having files. This included copying one of the existing files to use it for similar purposes or restoring one of the deleted files.

“I decided to just copy this video from my laptop to this hard drive.” (Pre-tenured Professor 1)

“I restored it about a few days ago because I had to print out more.” (Post-qual Graduate 1)

Save file in a temporary location. In the (3) Temporary Categorization stage, files were organized into a temporary category. In this case, participants delayed specific organizational decisions, but just saved an information item in easily accessible locations such as ‘Desktop’, ‘Downloads’, ‘Documents’, or ‘Dropbox’ instead of categories that fit into their folder structures. Thus, these categories were often used as a temporary work space with respect to that information object.

“I just saved it to my ‘Desktop’.” (Early Undergraduate 2)

“When you download something, it automatically put it in ‘Download’ folder.”
(Pre-tenured Professor 2)

Select a folder. In the (5) Selection/Modification/Creation stage, if participants have a relevant folder to categorize an information item, participants selected that folder from existing folders. For example, participants said:

“I put it in my ‘quiz #2’ folder.” (Late Undergraduate 2)

“I saved it in a folder that I already had.” (Late Undergraduate 3)

“I organized that file into ‘Conference name 2012 poster’.” (Pre-qual Graduate 1)

Modify previous folder. Sometimes, participants modified the category when they did not have an appropriate category to organize an information item, or when they thought some changes were necessary to better organize information files. There were three different ways of modifying existing categories. First, most often participants modified their existing categories by splitting the category. In this case, they created subordinate categories to split the category. Particularly, in the case of folders for classes, participants sometimes created a subfolder for each class he/she was taking in certain semester, a subfolder for each week of the class to organize readings for each week, or a subfolder for each exam or quiz scheduled in that class. Examples for each of them are as follows:

“I put the ones that have to do with one class into one folder. So I did that for a few of my classes.” (Early Undergraduate 1)

“I did make a week 2, 3, and 4 and moved all those stuff that are downloaded for those into them.” (Early Undergraduate 3)

“After I took quiz 2, I made a new folder...and then, I put all my lecture notes and the assignments before quiz 2 into here.” (Late Undergraduate 2)

Also, participants split their existing folders by different subtasks. For instance, one of the participants said:

“So, I took those files, and I split them up into the ones I was assigned to review.”

(Tenured Professor 1)

Second, sometimes participants modified their existing categories by merging folders. In this case, they created a superordinate category to group multiple categories. For instance, participants said:

“I actually created a new folder called ‘Travel picture’ folder and then I put it there. I put the folder into the bigger category.” (Pre-qual Graduate 3)

Often, participants mentioned that they will create a superordinate category to merge folders after they are done with a certain period of time, such as after a semester. For example, participants said:

“Usually, when I finish teaching, I move that folder into a macro-folder, which is called ‘Teaching and Advising’.” (Tenured Professor 3)

“I may just create a new folder, maybe just creating the bigger folder called ‘2012 Spring’ or ‘[2012] Summer’ or whatever and then just dump those files [folders] into that new folder.” (Pre-qual Graduate 1)

Third, participants modified their existing categories by changing the name of the categories. For example, participants said:

“I saved it in a folder that I already had, but you can see I changed the name of the folder.” (Late Undergraduate 3)

“It was just ‘ichat’? Yeah, because I started putting ‘ichat’, and I didn’t know what I would name another folder, just like pointless stuffs! So I just like added

‘ichat & such’ on there, so that it would encompass other things.” (Early Undergraduate 2)

Create a new folder. In addition, when participants could not find a folder appropriate to categorize an information item, they created a new category to organize an information item. For example, participants said:

“And then, I created a folder there.” (Tenured Professor 1)

“I created a new folder called ‘Project name’.” (Pre-qual Graduate 1)

As shown in Table 12, among 140 occurrences of actions shown in this stage (5), there were 109 occurrences of selecting one of the existing folders, 15 occurrences of modifying existing folders, and 16 occurrences of creating a new folder. This indicates that participants selected one of the existing folders more often than modifying existing folders or creating a new folder when they organize their information items. In addition, this shows that participants modified existing folders as frequently as they created a new folder.

Place file into a folder. In the (6) Categorization stage, participants placed the file into a folder which was one of the existing folders, modified folders, or newly created folders. For example, participants said:

“I just placed it there.” (Pre-tenured Professor 1)

“So I put this in this folder.” (Post-qual Graduate 3)

As shown above, although selecting a category and placing the file into a category are different actions, because these two actions happen continuously and quickly, it was hard to differentiate these two actions from the verbal descriptions from the participants.

However, participants did not show actions in all six stages. In the case of (2) Identification stage and (4) Examination/Comparison stage, it was difficult to identify any ‘actions’ that were involved in these stages. It is possible that in the (2) Identification stage, participants may check the name of the file, or check the creator, source, or time that are related to the file in an effort to identify the file. Also, participants may click files to open and see the content of the files to identify them. However, it was impossible to detect those behaviors from the diary or the interviews. Likewise, in the case of (4) Examination/Comparison stage, it is possible that participants may browse their existing folder structures or even click certain folders to examine existing categories. However, it was impossible to identify those behaviors from the diary or the interviews. It seems that perhaps an eye tracking method or computer data logging method can be used to elicit those data.

The behavioral aspect of the process of organizing personal information is presented in Figure 10.

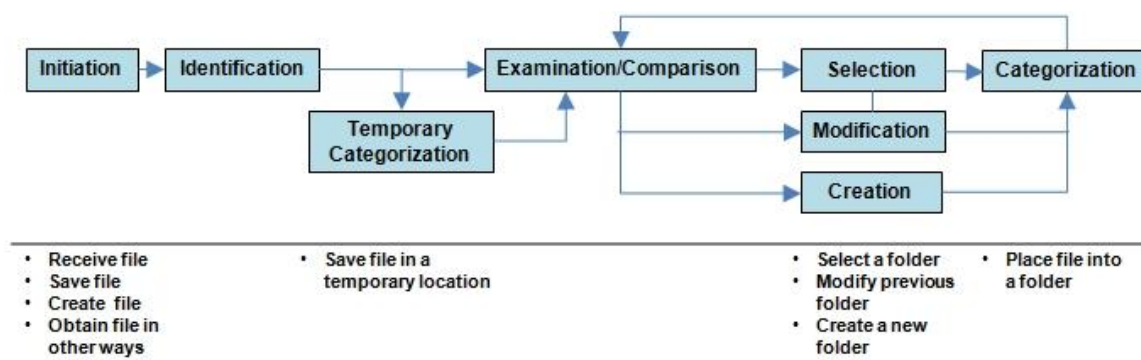


Figure 10. Behavioral Aspect of the Personal Information Organization Process

5.2.3 Cognitive Aspect

RQ3. What cognitive processes do people go through during the process of organizing personal information?

In exploring the process of organizing personal information, the cognitive aspect, which indicates any thought process participants went through or any feelings that were involved during the process of organizing personal information, has been identified and coded. Codes for thoughts that were developed while analyzing data and the occurrences of each code are organized by each stage of the process and presented in Table 13. Again, when counting occurrences, if the same code occurred multiple times in describing the process of organizing the same information item, they were counted as one. However, if different codes occurred in describing the process of organizing the same information item, each was counted separately.

Table 13
Codes for Cognitive Aspect of the Process

Stage	Code	Occurrence
(1) Initiation	Messy	0 ¹
	Confusing	2 ²
(2) Identification	Typify	93
(4) Examination/Comparison	Review existing folders	51
	Assess similarities and differences between new and existing files	52
	Adjust mental gap between new and existing files	67
(5) Selection/Modification/Creation		
(6) Categorization	Clean	0 ³

¹ This cognitive aspect was frequently mentioned by participants when they described about organization of information items which were not recorded in the diary because they were explaining about 1) their general organizing behaviors, 2) information items that were saved before they keep a diary for this study, or 3) what they will do in the future which were excluded when counting occurrences in the table.

² See footnote 1.

³ See footnote 1.

In the (1) Initiation stage, when information items are not organized, participants felt it was messy, confusing, and hard to find things. Table 13 shows that there were only few occurrences; however, this is because only those that were both (1) recorded in the diary, and (2) which have already taken place, were counted when counting occurrences. However, participants often used the expressions ‘messy’, ‘confusing’, and/or ‘hard to find things’ to describe their thoughts and feelings about the stage in which information items are unorganized when they were explaining their organizing behaviors or routines, describing information items that were organized before they kept the diary, or mentioning future or possible organizing processes. Although they were not counted as occurrences, because they are closely related to the organizing process of information items that were recorded in the diary, and because they are evidently part of their organizing behavior, these cognitive aspects of the process were incorporated into the personal information organization process model.

Messy. When information items are not organized, participants felt that ‘messiness’ became one of the motivations to initiate organization. For instance, when the researcher asked when they initiated organizing information items, participants said:

“At some point, when I think it’s really messy.” (Post-qual Graduate 1)

“I’m about ready to clean up my ‘Desktop’. It’s getting messy.” (Tenured Professor 1)

Also, when the researcher asked why they organized information items, participants mentioned ‘messiness’. For example, participants said:

“Because I was having a mess here.” (Post-qual Graduate 3)

“Things got so messy that I just threw them all in this folder.” (Tenured Professor 1)

Similarly, when participants described future or possible organization, they also mentioned ‘messiness’ they were having because information items were not organized. For instance, participants said:

“I want it all to be in ‘Midterm’ so that things don’t get all messed up.” (Post-qual Graduate 3)

“I think I should put the ‘Talks’ in by year, ’cause it’s pretty big and messy.”
(Tenured Professor 1)

This cognitive ‘messiness’ seems closely related to Douglas’ (1978) explanation about ‘dirt’. Douglas stated that ‘dirt’ is disorder, and people try to eliminate dirt in an effort to organize our environment (p. 2). Douglas also stated that the existence of ‘dirt’ means there is a system, as ‘dirt’ cannot exist alone (p. 35). In this study, it seemed that when there were information items that were not organized, which means that when information items are not part of the existing category structure, people felt it was messy, so they initiated an organization process in an effort to eliminate this ‘dirt’.

Confusing. When information items were not organized, participants not only felt messiness, but they also felt that it was confusing. For example, when the researcher asked why he/she organized information items, one of the participants said:

“I don’t like to get confused.” (Post-qual Graduate 3)

Also, one of the participants responded that he/she would organize information items so that it is not confusing. For instance, one of the participants said:

“I would probably create like a ‘Class name’ folder just for all my classes, and then within the folder like, I would like have sub-groups...so that it’s not confusing. (Early Undergraduate 1)

Similarly, participants used the expressions ‘hodgepodge’, ‘combined’, ‘mixed’, and ‘intertwined’ while describing how they felt when information items were not organized:

“Initially (before organizing) it was in a whole, hodgepodge whole collection of everything.” (Tenured Professor 1)

“I have things combined because these are two different projects, and like everything is mixed. So I decided to put things in different places.” (Post-qual Graduate 3)

“There were too many of them and it starts hard to see things because I got intertwined.” (Tenured Professor 1)

One of the participants also used the expression ‘irritated’ and ‘bothered’ when information items were not organized.

“There are times where like I’ll just suddenly get irritated like something’s not organized in a way that makes it easier for you to find something.” (Early Undergraduate 3)

“If it bothers me enough while I’m reviewing for the final, I might clean up and move them around.” (Early Undergraduate 3)

Typify file. In the (2) Identification stage, participants identified information items so that they could figure out how they should organize certain information items. In this process, ‘typifications’ occurred. Cooper (2004) stated that when we understand new information, we typify it according to previous experience and then put it into cognitive

categories (p. 302). Zerubavel (1991) also stated that “the ability to ignore the uniqueness of items and regard them as typical members of categories is a prerequisite for classifying any group of phenomena” (p. 17). In fact, participants typified information items by focusing on certain aspects of the object while disregarding other aspects of it. More specifically, when explaining how they organized a certain information item, they described that information item in ways that involved typifications. For example, participants said:

“It's a paper.” (Tenured Professor 4)

“This is from my father.” (Post-qual Graduate 1)

“It’s a picture.” (Pre-qual Graduate 3)

As shown above, although each information item was unique and had various aspects including purpose of the file, source of the file, format of the file, time related with the file, location related with the file, topic of the file, and so on, the participant typified an information item as ‘something’ by focusing on a certain aspect of the file while disregarding other aspects of it, so that they could categorize it into a certain category. In the first example, the participant focused on the purpose of the information item while overlooking other aspects of the file. Likewise, the participant in the second example focused on the source of an information item while ignoring other aspects. In the case of the third example, participant focused on the format of the information item. Sometimes, participants focused on more than one aspect of the file. For example, they said:

“This was PowerPoint slides for my other summer class.” (Late Undergraduate 3)

“This was also a PDF file from the web.” (Tenured Professor 3)

In the first example shown above, the participant focused on the format (“PowerPoint slides”) and purpose (“for my other summer class”) of the file, while the participant in the second example focused on the format (“PDF”) and source (“from the web”) of the file. Which aspect participants focus on when typifying information items is explained in more detail in Section 5.2.5.

In the (4) Examination/Comparison stage, there were two thought processes: 1) reviewing existing categories and 2) assessing similarities and differences between new and existing information items.

Review existing folders. In the (4) Examination/Comparison stage, to examine existing categories to see whether they have relevant categories to organize an unorganized item into, participants reviewed one or multiple categories they already had. For example participants said:

“I have another folder in ‘My Documents’ called ‘Reviews’, and embedded within that there’s a ‘Journal Name’ folder, I would just put it there.” (Tenured Professor 2)

“I’ve organized it by different research projects that I’m working on.” (Pre-tenured Professor 1)

“In that ‘Final analysis’ folder, I have all the analysis data, and the final paper.” (Post-qual Graduate 1)

“Everything related to my dissertation is here.” (Post-qual Graduate 2)

Thus, before making any organizational decision, participants reviewed relevant categories or their existing organizational structure.

Assess similarities and differences between new and existing files. In the (4) Examination/Comparison stage, participants also assessed similarities and differences between new and existing files in relevant categories so that they could decide the most appropriate category to organize a new information item. For instance, participants said:

“I could put it under ‘Reviews’ [folder] but, it’s really not quite reviews, it’s the assignments of what I’m supposed to read, and what everybody else is supposed to read.” (Tenured Professor 1)

“It’s not technically class work, but it is part of the summer work that I am doing for school this semester.” (Late Undergraduate 3)

“So it’s a conference thing, but it’s not really a publication or a paper. As I say, it’s a tutorial.” (Pre-tenured Professor 1)

Here, in the first example, the participant looked at one of his/her folders, which is ‘Reviews’ folder, and then described the differences (“it’s really not quite reviews”) between an unorganized information item and organized information items in that folder. In the second example, the participant assessed differences (“it’s not technically class work”) as well as similarities (“but it is part of the summer work that I am doing for school this semester”) before deciding where to categorize an unorganized information item. Similarly, in the third example, the participant assessed similarities (“it’s a conference thing”) and differences (“but it’s not really a publication or a paper”) between new and existing information items in a relevant category. Thus, as shown in these examples, when participants reviewed a relevant category or categories to organize an information item, they assessed similarities and differences between new and existing files in those relevant categories.

Adjust mental gap between new and existing files. In (5)

Selection/Modification/Creation stage, participants adjusted their mental gap between new and existing information items, which resulted in selecting or modifying one of the existing categories, or creating a new category. When participants selected one of the existing categories, they adjusted their mental gap between unorganized and organized information items by exaggerating the similarities between the new file and existing files in that category and reducing the differences between them. Zerubavel (1991) also stated that when people categorize things, they “let their similarity outweigh any differences among them” (p. 16). For example, participants said:

“It’s a review, so I put it under ‘Reviews’.” (Tenured Professor 1)

“I saved it there because it was taken on the same day as the other pictures, and it has to do with them.” (Late Undergraduate 2)

“I was deciding if I should make a new folder just for the review questions? But then I decided that it would probably be easier, more practical, if I added it to folder that I already have all the information about the class, so that it would be easier to like, relate back to.” (Early Undergraduate 1)

Here, to select one of the existing categories, participants adjusted their mental gaps by reducing the differences between new and existing information items and emphasizing the similarities between them. In the first example, although the file is typified as a ‘review’, it may have various aspects that are different from existing information items that are organized in ‘Reviews’ category in terms of its format, time, source, topic, main use, related person, and so on. However, the participant set aside the differences between them so that this new information item could be part of the information items in the

‘Reviews’ category. At the same time, the participant emphasized the similarity between new and existing information items, which is they are ‘review’ of something or someone so that they can be grouped together. In a similar vein, in the other examples, participants exaggerated the similarities between new and existing information items in a selected folder, such as time (“it was taken on the same day”), format (“pictures”), topic (“it has to do with them”), and purpose (“information about the class”) of the information items while ignoring the differences between new and existing information items. At the same time, selecting a certain category means that participants exaggerated the differences and reduced the similarities between new information items and information items in the unselected categories. For example, when one of the participants selected a category, he/she said:

“I put it in my ‘quiz #2’ folder, because that was an assignment, that was assigned to me during the time where, during the time after quiz #1 and then before quiz #3.” (Late Undergraduate 2)

Here, although new information items have various aspects that are similar to the information items in the relevant categories (i.e., ‘quiz #1’ and ‘quiz #2’ folder), the participant deselected ‘quiz #1’ and ‘quiz #3’ categories, because the time new information item was assigned to him/her was different from information items in ‘quiz #1’ and ‘quiz #3’ categories.

When participants failed to adjust the mental gap between new and existing information items in the existing category by inflating the similarities and deflating the differences to select one of the existing categories, participants modified the existing

category so that they could successfully adjust their mental gaps between them, and categorize the new information item. For example, participants said:

“It was just ‘ichat’? Yeah, because I started putting ‘ichat’... but then I changed it to ‘ichat & such’ because I’ll just put any other, like pictures, funny stuff that people send me whatever went into there, too.” (Early Undergraduate 2)

“I mentioned to you that I [will] create subfolders within that folder. I’ll have one for each of my three classes, so I’m going to do that this weekend and then sort the files accordingly.” (Tenured Professor 3)

In the examples above, participants made some changes to the existing categories to successfully adjust mental gaps between new and existing information items. In the first example, the participant originally had an ‘ichat’ folder where he/she saved screen shots that he/she took during a video chat with his/her friends. However, when saving a funny picture, although he/she found similarities between that picture with pictures that were saved in the ‘ichat’ folder, i.e., random fun pictures, he/she failed to inflate similarities enough to select a ‘ichat’ folder because the category was limited to pictures from chatting, which is a difference between new and existing information items he/she could successfully reduce or ignore. Thus, he/she modified the name of the folder to ‘ichat & such’ so that he/she could successfully adjust the mental gap by exaggerating the similarities and overlooking differences between new and existing information items so that he/she could successfully categorize the new information item.

When participants could not adjust their mental gaps between new and existing information items by modifying existing categories, they created a new category. In these cases, participants exaggerated the difference between an unorganized information item

and information items that were already categorized in existing categories. For instance, participants said:

“I don’t really have a folder that would have been fitting specifically for this.”

(Early Undergraduate 3)

“It’s a totally new job, so I don’t want to conflict with my existing folders.” (Pre-qual Graduate 1)

Here, although there might be some similarities between new and existing information items in terms of time, format, purpose, source, and so on, the participants ignored those similarities and exaggerated the differences, while saying that he/she did not have a category for that information item. Then, participants created a new category to organize new information items.

Clean. In the (6) Categorization stage, in contrast to the ‘messiness’ participants cognitively felt when information items were not organized, participants felt it was clean when an information item was organized. However, just like ‘messy’ and ‘confusing’ codes in the (1) Initiation stage, participants did not actually use the expression ‘clean’ while describing the organizing process of information items that were recorded in the diary. However, while giving explanations about information items that were organized before they kept a diary, or future or possible organizing processes, they often used the expression ‘clean’ to describe their thoughts and feelings about the stage in which information items were organized. Thus, the researcher included it as a cognitive aspect of the personal information organization process. For example, participants used the expression ‘clean’ when they described possible or future personal information organization. For example, participants said:

“It will eventually be cleaned up and go into the folder.” (Tenured Professor 1)

“Usually like once a month I’ll go through and clean everything up so I didn’t really get to it, I didn’t know if I would still need it with this project. So, probably after our meeting tonight this weekend, I’ll probably put it in another folder.”

(Late Undergraduate 3)

“If it bothers me enough while I’m reviewing for the final, I might clean up and move them around.” (Early Undergraduate 3)

“At the end of each year I’ll do a cleanup of the ‘Desktop’, move stuff off.”

(Tenured Professor 2)

Similarly, one of the participants used the term “clear” as follows:

“I would move stuff the end of the summer and clear the ‘Desktop’ for fall material.” (Tenured Professor 2)

Participants also felt that it was “simple” when information items are organized. For example, when the researcher asked a participant why he/she grouped certain information items together, the participant said:

“To make it simpler, to make my folder look simpler and collect similar things in the same category.” (Post-qual Graduate 3)

Thus, when information was placed into a category, participants felt it was cognitively clean, clear, and simple.

However, not all stages involved cognitive processes. In addition, as emphasized previously, people did not always go through certain cognitive processes in any particular stage of the model. Therefore, what is presented in this model only shows ranges of

thoughts that are involved during the process of organizing personal information, rather than typifications of thoughts.

Cognitive aspects of the process of organizing personal information are presented in Figure 11.

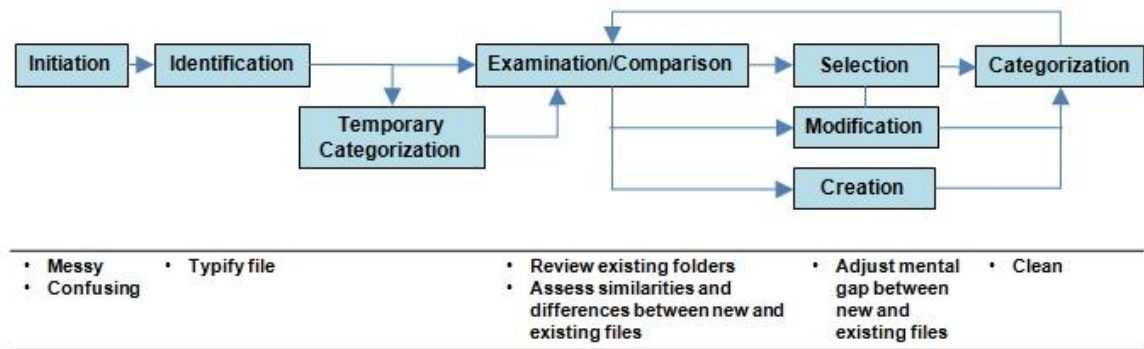


Figure 11. Cognitive Aspect of the Personal Information Organization Process

5.2.4 Decisions

RQ4. What decisions do people make during the process of organizing personal information?

There were various decisions involved during the process of organizing personal information. While analyzing data, any organizational decisions made during the process were identified and coded. Codes for decisions developed during analyzing data as well as occurrences of each code are organized by each stage of the process, and presented in Table 14.

Table 14
Codes for Decisions Made During the Process

Stage	Code	Occurrence
(1) Initiation	Save file	113
(3) Temporary Categorization	Delay decision making	23
	Re-categorize file	18
	Delete file	13
	Keep temporary categorization	8
(5) Selection/Modification/Creation	Select a folder	99
	Modify previous folder	15
	Create a new folder	14
(6) Categorization	Delay decision making	41
	Re-categorize file	15
	Delete file	8
	Keep categorization	83
	Move file	3

Save file. In the (1) Initiation stage, the first decision participants made was whether to save an information item or not. This decision included deciding to save information items in a temporary location, in the existing folders, or in a new folder. For example, participants said:

“I just saved that file into ‘Conference name 2012 poster’ folder.” (Pre-qual Graduate 1)

“Then I decided to save it.” (Pre-qual Graduate 2)

“So I just saved it to the Download folder.” (Pre-tenured Professor 1)

Delay decision making. The third stage of the process of organizing personal information is called (3) Temporary Categorization, because in this stage, participants often delayed making any organizational decisions. The analysis of the data showed that among decisions participants made in the (3) Temporary Categorization stage, ‘delaying decision making’ occurred most frequently (37.1%). During the interview, participants said:

“I dump everything on ‘Desktop’, and then I file later.” (Tenured Professor 3)

“‘Downloads’ is just a temporary place where anything from a web browser or email client gets saved temporarily, and then I make a decision if I want to delete it, move it.” (Pre-tenured Professor 1)

In the case of delayed decisions, there were two different types. One type of delayed decision is when participants know what decisions they are going to make in the future but delay making decisions for the time being. In these cases, after saving information items in a temporary location, participants said:

“I’ll put it in a folder for the summer session 2012.” (Late Undergraduate 3)

“After this exam, I can delete that PowerPoint, and never worry about it, so there was no point in saving it.” (Late Undergraduate 1)

In these examples, participants knew about delayed decisions. In the first example, the participant knew that he/she would re-categorize an information item into a specific folder (“a folder for the summer session 2012”). In the case of second example, the participant knew that he/she would delete an information item in the future (“delete that PowerPoint”). The other type of delayed decision is when participants not only delay making decisions but also are not certain about specifically what decisions they are going to make in the future. In these cases, participants said:

“I haven’t decided where to put, I put it on the ‘Desktop’, when I clean it, I can just put it into the right folder.” (Pre-tenured Professor 2)

“Either I’ll move it right away, or I’ll just start editing the file, save it, and then move it.” (Pre-tenured Professor 1)

In the first example, the participant knew that he/she would re-categorize it (“put it into the right folder”); however, he/she did not exactly decide in which folder he/she would organize the information item (“haven’t decided where to put”). In the second example, the participant did not even know whether he/she would re-categorize the information item (“move it right away”) or keep it for a while (“start editing the file, save it, and then move it”). In this research study, among 23 delayed decisions, 21 decisions (91.3%) belonged to the first type while only two decisions (8.7%) belonged to the latter. Thus, when delaying decisions, most of the time, participants knew what they were going to do with the information item in the future, but simply delayed making decisions temporarily.

The further analysis of the data showed that among 21 decisions for which participants knew specifically what decisions they were going to make in the future, 14 of them were ‘delaying re-categorization’ (66.7%) while 6 of them were ‘delaying deletion’ (28.6%). In addition, one of them was ‘delaying moving’ to another device (4.8%). Thus, most of delayed decisions were re-categorization. In addition, in all 14 delayed decisions of re-categorization, participants decided to organize information items into one of the categories in their folder structures. For example, participants said:

“And all of these files will go in that folder (‘Fall Courses’ folder). I just haven’t filed them yet.” (Tenured Professor 3)

“I saved it to my ‘Desktop’, which I will move into my ‘Spring Semester’ folder, and into the subfolder that I have in my ‘Spring Semester’ folder called ‘Course 1’, but I haven’t done that yet.” (Late Undergraduate 1)

Thus, in the (3) Temporary Categorization stage, participants mostly knew what decisions they were going to make, and those delayed decisions were primarily re-categorizing information items into one of the existing folders in their folder structures.

Re-categorize file. As previously described, most of the ‘delayed decisions’ in the (3) Temporary Categorization stage were ‘re-categorization’. As a matter of fact, the second interview, which was conducted 2-4 weeks after the first interview, showed that participants actually did re-categorize information items that were saved in the temporary location. In the (3) Temporary Categorization stage, among 58 decisions, 18 of the decisions were ‘re-categorization’ (31.0%), which is the second highest number of occurrences. In these cases, when the researcher asked participants what happened to the information items that were saved in the temporary locations, participants said:

“I transferred to my ‘iPhoto’, which I then put in the ‘St. Patty’s Day 2012’ folder.” (Late Undergraduate 1)

“That’s also in the subfolder, now.” (Early Undergraduate 1)

“I first downloaded on my ‘Desktop’. And then I decided what to do with it and then created a folder.” (Pre-qual Graduate 3)

Here, in the first example, the participant re-categorized an information item into one of the existing folders (“I then put in the ‘St. Patty’s Day 2012’ folder”). In the second example, the participant split the category by creating subfolders. In the third example, the participant re-categorized an information item by creating a new category (“created a folder”).

Sometimes participants re-categorized information items not only into the folder structure in the same personal device but also into another personal device. For example, one of the participants said:

“I put it on my external hard drive into my ‘Spring Semester’ folder.” (Late Undergraduate 1)

Here, the participant first moved the information item into another personal device (“I put it on my external hard drive”) and then organized it into one of the existing categories (“into my ‘Spring Semester’ folder”).

The further analysis showed that when re-categorizing, 16 decisions out of 18 were categorizing information items into one of the existing categories (88.9%), while one decision was modifying a category (5.6%), and another one decision was creating a new category (5.6%). Thus, similar to the delayed re-categorization decisions, when participants re-categorized information items that were saved in a temporary location, most of the time participants organized them into their existing organizational structures. This result shows that (3) Temporary Categorization stage is indeed different from (6) Categorization stage.

Delete file. In the (3) Temporary Categorization stage, often information items were deleted. When asked questions about if there had been any changes made to the information items that were discussed in the first interview, participants said:

“I erased it.” (Early Undergraduate 2)

“I put it in the recycle bin.” (Pre-qual Graduate 3)

“I threw them away.” (Post-qual Graduate 1)

This decision was understandable because some of the characteristics of information items that were saved in the temporary locations included information items that were needed only for a short period of time or information items that were not important. The factors that impact on (3) Temporary Categorization stage are discussed in more detail in Section 5.2.5.

Keep temporary categorization. As stated previously in Section 5.2.1, in the (3) Temporary Categorization stage, participants saved information items in a temporary location for only a limited time. However, the length of this ‘temporary’ time varied greatly. More specifically, sometimes it was as short as a few minutes as shown in the following example:

“I saved it to my ‘Desktop’, and then I had erased it like right after I send it to my friend.” (Early Undergraduate 2)

However, sometimes, it was as long as a year, as stated in the following example:

“At the end of each year I’ll do a cleanup of the ‘Desktop’, move stuff off.”
(Tenured Professor 2)

Thus, during the 2-4 weeks between the first and the second interview, some information items were kept in the temporary locations, which accounted for 13.8% of the decision occurrences. For example, participants said:

“It’s still there.” (Pre-tenured Professor 2)

“It’s still on my ‘Desktop’. And I haven’t done the work.” (Tenured Professor 1)

Thus, although these information items would be re-categorized or deleted sometime in the future, for the time being some of the information items were kept in the temporary location.

Select a folder. In the (4) Selection/Modification/Creation stage, participants decided to select, modify, or create a folder. In this stage, among 128 decision occurrences, 99 of them were ‘selecting a folder’ (77.3%). This indicates that in the fourth stage of the personal information organization process, participants primarily decided to select one of the existing categories. During the interview, while describing how they organized their personal information items, participants often said that they saved them in one of the existing folders. For example:

“[I saved it] in the laptop, and existing folder called ‘Course 1.’” (Pre-qual Graduate 1)

“I saved it under ‘Teaching’.” (Pre-tenured 1)

“That goes under ‘Teaching’ and then the ‘Class’ and then in ‘Slides’.” (Post-qual Graduate 2)

Modify previous folder. Sometimes, participants decided to modify one of the existing folders in the (4) Selection/Modification/Categorization stage. Among 128 decision occurrences in the fourth stage of the personal information organization process, 15 (11.7%) were deciding to modify an existing folder. This modification decision included deciding to (1) split a category by creating subordinate categories, (2) merge categories by creating a superordinate category, and (3) change the name of the category. Examples for each case are as follows:

“I decided to split the folders into weeks as well. So now in each of the classroom folders, I have subfolders for ‘Week 1’, ‘Week 2’, ‘Week 3’, and ‘Week 4’.”

(Early Undergraduate 3)

“So at some point, I decided to just have one ‘Conference name’ folder. And under ‘Conference name’ folder, I have subfolders, which are ‘2010’, ‘2011’, that way.” (Pre-tenured Professor 1)

“I saved it in a folder that I already had, but you can see I changed the name of the folder” (Late Undergraduate 3)

Here, in the first example, the participant decided to split the folder (“I decided to split the folders into weeks”) while the second participant decided to merge categories by creating a superordinate folder (“I decided to just have one ‘Conference name’ folder”). In the last example, the participant decided to change the name of the folder (“I changed the name of the folder”).

Create a new folder. In the (4) Selection/Modification/Creation stage, participants also decided to create a new folder to organize information items. Among 128 decision occurrences in the fourth stage of the personal information organization process, 14 (10.9%) were creating a new folder. For instance, participants said:

“I created a new folder called ‘Project name’.” (Pre-qual Graduate 1)

“I’ve made a new folder. And it’s called ‘University 1’.” (Post-qual Graduate 2)

In the (6) Categorization stage, decisions were investigated through the second interview. Thus, it is important to note that all of these decisions were made during the 2-4 weeks between the first and the second interview.

Delay decision making. In the (6) Categorization stage, sometimes participants delayed making decisions, just as they did in the (3) Temporary Categorization stage. In this stage, delayed decisions included delaying re-categorization of information items,

delaying deleting information items, and delaying moving information items into other personal devices. For example, participants said:

“When I finish these courses in the fall, I’ll put them into that ‘Teaching’ folder under ‘Course 1’.” (Tenure Professor 3)

“I will delete it when I’m done with everything.” (Tenured Professor 1)

“When the next semester starts, then, I’ll take this, everything, like all this ‘spring 2012’, I’ll move it to external drive, as a whole.” (Late Undergraduate 2)

In the first example shown above, the participant delayed re-categorizing an information item to another category (“I’ll put them into that ‘Teaching’ folder under ‘Course 1’”). In the second example, the participant delayed deleting an information item (“I will delete it”), and in the third example, the participant delayed moving information item to another personal device (“I’ll move it to external drive”).

This process looks similar to ‘delay decision making’ in (3) Temporary Categorization. However, there were some differences between delaying decisions in the (3) Temporary Categorization stage and the (6) Categorization stage. First of all, in the (3) Temporary Categorization stage, ‘delaying decision making’ was the decision with the highest occurrence. However, it did not show the highest occurrence in the (6) Categorization stage. In addition, in the (3) Temporary Categorization stage, among 21 decisions for which participants knew specifically what decisions they were going to make in the future, 14 of them were ‘re-categorization’ (66.7%) while 6 of them were ‘deletion’ (28.6%), and one of them was ‘move’ (4.8%). However, in the (6) Categorization stage, among 41 delayed decisions, which participants knew what decisions they delayed, 28 of them were ‘re-categorization (68.3%) while 8 of them were

‘deletion’ (19.5%), and 5 of them were ‘move’ (12.2%). Thus, there was a greater percentage of information items that were delayed to be deleted in the (3) Temporary Categorization stage than in the (6) Categorization stage. Most importantly,, in the (3) Temporary Categorization stage, in the case of 14 delayed re-categorization decisions, all of them (100%) were delayed to be organized into one of the categories in the folder structures. In the (6) Categorization stage, however, among 28 delayed re-categorization decisions, 20 of them (71.4%) were moved into one of the existing folders, while 8 of them (28.6%) were delayed to be re-categorized by modifying existing categories, such as splitting or merging categories. This analysis shows that while delayed decisions were simply delaying organizing information items into the existing categories in the (3) Temporary Categorization stage, the delayed decisions in the (6) Categorization stage involved changes to the existing category structure.

Re-categorize file. In the (6) Categorization stage, sometimes information items were re-categorized. Among 150 decision occurrences in the (6) Categorization stage, 15 (10.0%) of them were re-categorization. These re-categorization decisions that were found during the second interview included re-categorizing into another folder and modifying folders in which information items were originally saved by creating subordinate folders or superordinate folders. For example, participants said:

“I fixed those. That’s under ‘Travel Advice’. And I put all those letters there.”

(Tenured Professor 1)

“I made an ‘Exam 2’ folder, and then I put the chapter 9 notes and that.” (Late

Undergraduate 2)

“I actually created a new folder called ‘Travel picture’ folder and then I put it there. I put the folder into the bigger category.” (Pre-qual Graduate 3)

Here, in the first example, the participant re-categorized an already organized information item into a different folder. In the second example, the participant modified a folder in which he/she organized an information item by splitting it. In the third example, the participant modified the folder where he/she originally categorized an information item by merging similar folders into a bigger folder.

Delete file. Sometimes, participants deleted information items that were categorized into a folder in the (6) Categorization stage. The second interview showed that among 150 decision occurrences, 8 decisions (5.3%) were deleting files. This indicates that during 2-4 weeks between the first and the second interview, eight files were deleted. For example, participants said:

“I erased that folder, actually the whole folder.” (Early Undergraduate 2)

“I actually deleted it.” (Late Undergraduate 1)

Keep categorization. In the (6) Categorization stage, the ‘keeping categorization’ decision represented more than half of the decision occurrences (55.3%). That is, participants decided to keep the information items in the folder where they initially organized them. For example, participants said:

“It’s still there.” (Pre-qual Graduate 2)

“That’s so same, I didn’t move.” (Early Undergraduate 2)

“They are still in the same folder.” (Post-qual Graduate 3)

This result shows that as much as information items are re-categorized (10.7%), deleted (5.3%), moved (2.0%), or delayed to be changed (26.7%), they were also kept in a folder.

However, in this stage, many participants said that there were possibilities of making changes to these information items, so that 30 different information items were shown to have possibilities of getting changed. When identifying possible decisions, the researcher examined the answers to the question, “Q7: Do you think there is a possibility of moving this information item into other place sometime later?”, which was the last question of the first interview. In these cases, unlike the delayed decisions in which participants often used the expression ‘will’, participants used the weaker expressions in terms of their willingness by using the expressions such as ‘may’, ‘maybe’, ‘might’, ‘could’, or ‘possibly’. These possible future decisions included re-categorizing, deleting, and moving. Examples are as follows:

“Maybe I can reorganize those folders according to the format in addition to the year and the conference.” (Pre-tenured Professor 1)

“But at some point we might want to, I might want to do something else with it and move it someplace else. So that’s something that could get moved potentially to another folder or file.” (Post-qual Graduate 2)

“I could back it up and then transfer all these files to another computer.” (Tenure Professor 4)

Here, in the first example the participant mentioned the possibility of having a different organizational structure (“maybe I can reorganize those folders according to the format”). In the second example, the participant described the possibility of re-categorizing an information item into a different category (“that could get moved potentially to another folder”). In the last example, the participant mentioned the possibility of moving information items into another personal device (“transfer all these

files to another computer”). This result shows that it is almost impossible to have a perfect and complete organizational structure for personal information. Rather, it keeps changing over time, although the length of time it takes for any changes to occur varies. However, in the (3) Temporary Categorization stage, the second interview showed that during 2-4 weeks between the first and the second interview, there were 8 occurrences of ‘keeping temporary categorization’ (13.8%), which shows that only a few of the information items were kept without any change. In the case of the (6) Categorization stage, there were 83 occurrences (55.3%) of ‘keeping categorization’ decision, which was much higher than for the (3) Temporary Categorization stage. This result shows that it usually takes a longer time for any changes to occur to the information items in the (6) Categorization stage when compared to information items in the (3) Temporary Categorization stage.

Move file. Although there were just few occurrences (2.0%) at the time of the second interview, sometimes information items were not only kept in a folder but also got moved to participants’ other personal devices such as external hard drives or web sites. For example, participants said:

“I decided to just copy this video from my laptop to this hard drive.” (Pre-tenured Professor 1)

“Those I put on the online class management system.” (Tenured Professor 4)

As shown above, there were various decisions involved during the process of organizing personal information. However, not all six stages of the process involved decision making. The researcher could not identify decisions that were made in the (2) Identification stage or (4) Examination/Comparison stage. Interestingly, these two stages

were the stages in which the researcher could not identify ‘actions’ either. Thus, it seemed that one of the possible reasons for the difficulty the researcher had in identifying decisions made in these stages was because the researcher could not identify actions taken in these stages. In other words, because actions are the results and the evidence of decision making, it was difficult to identify decisions when the researcher could not identify actions.

Decisions that are made during the process of organizing personal information are presented in Figure 12.

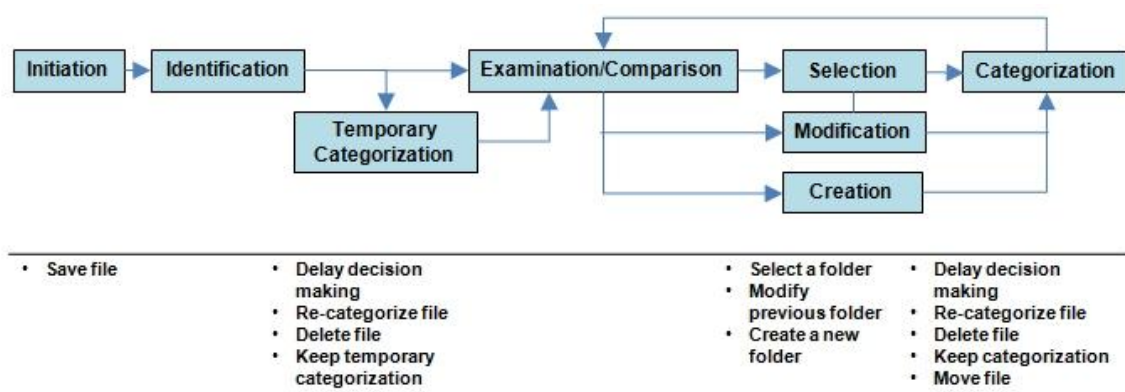


Figure 12. Decisions Made During the Personal Information Organization Process

5.2.5 Factors

RQ5. What factors influence the process of organizing information?

There were various factors that affected the process of organizing personal information. Investigating what factors influence the personal information organization process helps to understand why participants took certain actions, went through certain thought processes, and/or made certain decisions. Thus, any factors that influenced actions, thoughts, or decisions in each stage of the process were identified and coded

while analyzing data. To investigate possible factors as much as possible, when counting codes, not only factors that influenced any actions, thoughts, and decisions that took place or were made to the information that was kept in the diary, but also factors that impacted on any possible decisions that were mentioned by participants were taken into account. However, like code occurrences for actions, thoughts, and decisions, any mentions of organizing behavior of information items that were not recorded in the diary were excluded when counting occurrences. Codes for factors, frequency of overall occurrences, as well as frequency of occurrence in each stage are presented in Table 15.

Table 15
Codes for Factors Impact on the Process

Code	Overall Occurrence	Occurrence by Stage	
Accessibility	75	(1) Initiation	3
		(3) Temporary Categorization	18
		(4) Examination/Comparison	2
		(5) Selection/Modification/Creation	24
		(6) Categorization	28
Affiliation	14	(4) Examination/Comparison	5
		(5) Selection/Modification/Creation	9
Appropriateness	12	(6) Categorization	12
Availability	7	(3) Temporary Categorization	2
		(5) Selection/Modification/Creation	5
Format	41	(2) Identification	26
		(4) Examination/Comparison	9
		(5) Selection/Modification/Creation	6
Messiness	10	(3) Temporary Categorization	4
		(6) Categorization	6
Necessity of differentiation	23	(5) Selection/Modification/Creation	14
		(6) Categorization	9
Number of files	25	(5) Selection/Modification/Creation	13
		(6) Categorization	12
Purpose	250	(2) Identification	104
		(4) Examination/Comparison	47
		(5) Selection/Modification/Creation	88
		(6) Categorization	11
Related person	42	(2) Identification	24
		(4) Examination/Comparison	10

		(5) Selection/Modification/Creation	8
Reminder	9	(3) Temporary Categorization	6
		(5) Selection/Modification/Creation	3
Source	24	(2) Identification	24
System	7	(1) Initiation	2
		(3) Temporary Categorization	2
		(5) Selection/Modification/Creation	3
Time	66	(2) Identification	25
		(4) Examination/Comparison	18
		(5) Selection/Modification/Creation	23
Time availability	19	(3) Temporary Categorization	11
		(5) Selection/Modification/Creation	2
		(6) Categorization	6
Topic	47	(2) Identification	26
		(4) Examination/Comparison	7
		(5) Selection/Modification/Creation	14
Type	34	(2) Identification	28
		(4) Examination/Comparison	6
Use	175	(1) Initiation	36
		(2) Identification	15
		(3) Temporary Categorization	36
		(4) Examination/Comparison	2
		(5) Selection/Modification/Creation	14
		(6) Categorization	72
Value	20	(1) Initiation	4
		(2) Identification	6
		(3) Temporary Categorization	3
		(5) Selection/Modification/Creation	2
		(6) Categorization	5

Accessibility. One of the factors that influenced the process of organizing personal information was the ‘accessibility’ of an information item. The ‘accessibility’ factor was one of the influential factors, in that it represented the third highest number of occurrences among 19 factors. In addition, this factor affected almost all stages of the process except the (2) Identification stage.

In the (1) Initiation stage, participants decided to save information items to facilitate access to the information items. For example, when information items were on

web sites, participants downloaded them so that they could easily access them. The analysis of the results showed that downloading information items from web sites not only made it easier for the participants to access information items, but it was also more stable because participants did not have to worry about internet access. For example, participants said:

“I save like the different reports from the different units, so that I have access to them whenever we’re having a discussion that they were all available to me. I could go to web site, but that’s kind of contingent upon the University Wireless network, which sometimes doesn’t work very well. It depends on what building you’re in as well.” (Tenured Professor 2)

“Because the University web site platform is not very easy to use. It’s very slow, and a couple of times it actually wouldn’t let me in that day. So I got tired of trying to look up and try to figure things out. I ended up saving it to make it easier for me to refer to by putting in my own personal storage.” (Early Undergraduate 3)

Here, in the first example, the participant mentioned saving information items to facilitate access (“so that I have access to them whenever we’re having a discussion that they were all available to me”), as well as instability of the web site (“sometimes doesn’t work very well”). In the second example, the participant explained about saving information items for easy access (“make it easier for me to refer to”), and the inconvenience and instability of web platform he/she had to use (“It’s very slow, and a couple of times it actually wouldn’t let me in”).

In the (3) Temporary Categorization stage, accessibility of information item also affected the decision of saving files into temporary locations. Participants often decided to save an information item into a temporary location when the files needed to be accessed quickly, frequently, or immediately. For example, when the researcher asked why they saved certain information items into temporary locations, participants said:

“I just figured I’d be accessing it quickly if someone needed me to send it to them.” (Late Undergraduate 3)

“To view them more frequently and conveniently.” (Pre-qual Graduate 3)

“I saved it to ‘Desktop’ because I knew I would need it quickly.” (Tenured Professor 3)

“I thought I would have to be accessing it more, so it’s easier instead of having to go through a folder.” (Late Undergraduate 3)

Especially, the last example shows that saving an information item in a temporary location enhances accessibility more than saving it in one of the folders in the folder structure because participants do not have to go through the folder. As a matter of fact, among 75 occurrences of ‘accessibility’ code, 18 of them (24%) impacted on the (3) Temporary Categorization stage.

In the case of (4) Examination/Comparison stage, among 75 occurrences of the ‘accessibility’ code, there were only two occurrences (2.7%) that influenced this stage. In these cases, when participants reviewed existing folders and compared unorganized and organized information items in relevant folders to decide where to save an information item, ‘accessibility’ was one of the factors they considered. For example, while

describing about how he/she examined existing folders and compared new and old information items to decide where to save an information item, a participant said:

“I could have done it in there (save in 'Documents' folder), too. I think I just looked in the ‘ichat & such’ folder more than I look in my regular ‘Documents’, because I usually don’t save things there, so I wouldn’t go looking for things in there probably.” (Early Undergraduate 2)

In this example, the participant reviewed both ‘Documents’ and ‘ichat & such’ folders, and when comparing an unorganized information item with organized information items in those folders to decide which folder to place an unorganized information item, ‘accessibility’ was one of the factors that influenced the decision (“I just looked in the ‘ichat & such’ folder more than I look in my regular ‘Documents’”).

In the case of (5) Selection/Modification/Creation stage, 24 of the 75 occurrences of ‘accessibility’ impacted on this stage, which took 32% of all ‘accessibility’ code occurrences. In this stage, accessibility of an information item influenced participants’ decision making in two different ways. The first way was when participants saved information items into a folder in temporary locations. This is different from those that were saved in temporary locations without folder or in a default folder such as ‘Downloads’ or ‘Documents’. However, the reason why participants saved information items into a folder that was located in the temporary place was heavily influenced by an ‘accessibility’ factor. For example, one of the participants who saved an information item into a folder called ‘2012 Work’ that was located in the ‘Desktop’ said:

“It’s what I’m working on immediately. It’s available to me on the ‘Desktop’.
(Tenured Professor 2)

The second way that accessibility of an information item affected the decision of selecting a category to organize an information item was when participants saved information items into one of the categories in their folder structures so that they could easily find them. In these cases, participants needed to go through multiple levels of folder structures, and whether participants could easily remember where they saved information items, and whether they could easily find them influenced the decision. Thus, when the researcher asked why they saved an information item into a certain category, participants said:

“That way I can see exactly where the midterm (file) is. I don’t have to hunt for it.” (Post-qual Graduate 2)

“That way I could just, you know, easily find it.” (Early Undergraduate 1)

“I have it embedded in this folder, so that when I do get to the point of actually submitting a conference proposal, I know where to find the description of the conference.” (Tenured Professor 2)

As shown in these examples, in these cases, rather than immediate and quick access, findability of an information item influenced the decision. In addition, this ‘findability’ was closely related to ‘memory’ so that it was easier for participants to find certain information items when they could readily remember where they saved them.

Similarly, in the (6) Categorization stage, the ‘findability’ of information items affected the decision. During the second interview, when the researcher asked participants why they decided to keep certain information items into a category where they saved it, participants said:

“It’s the easiest place for me to keep them so I don’t forget.” (Post-qual Graduate 2)

“That’s just exactly where I would look for it if I needed to get it again.” (Pre-qual Graduate 2)

In addition, in this stage, participants decided to keep categorization when they could easily access information items within a category. One of the participants said:

“At the end of that trail, I have maybe 10 objects in that folder, which is fine. I can sort through that. I don’t need to search them. I can quickly read through those names and immediately know that okay, that one is a paper. That one is a poster, so that’s fine.” (Pre-tenured Professor 1)

However, the same participant said:

“But if I had 40 different things in there, then I would probably want to break it down under more subfolders”. (Pre-tenured Professor 1)

Thus, participants decided to keep categorization in the (6) Categorization stage when they could easily find needed information items.

In addition, in the (6) Categorization stage, participants sometimes decided to move an information item into another personal device because of accessibility. For instance, participants said:

“I decided to just copy this video from my laptop to this hard drive. And then I could access that hard drive from anywhere.” (Pre-tenured Professor 1)

Affiliation. In the process of organizing personal information, affiliations of participants, which are groups or institutions to which participants are formally connected or joined (such as participants’ university or program) influenced the decisions.

This factor affected the (4) Examination/Comparison stage as well as (5) Selection/Modification/Creation stage. When participants examined existing folders while assessing similarities and differences between an unorganized item and organized information items in relevant folders, or when they finally adjusted a mental gap between new and existing information items, the ‘affiliation’ relevant to an information item affected the process. For example, when the researcher asked why participants saved an information item in a certain category, participants said:

“It’s a folder that I created after I came to University 1, so basically, everything that I needed to store related to University 1 is here.” (Pre-qual Graduate 3)

“I know before like 2010, probably I can find the things from here (‘University 1’ folder), and then after I came to University 2, everything will be here.” (Pre-tenured Professor 2)

“Because they are related to my job as being a professor at University 1.”
(Tenured Professor 1)

As shown in these examples, participants often created a category by the name of the formal affiliation, and saved information items that were created or saved while they were a member of that affiliation. Thus, if their affiliations changed, they created another category based on new affiliation name, and started to save information items to that category. In addition, even after their formal affiliations had been changed, if they obtained information items that were relevant to any work or person in that affiliation, they were also saved into that category. Thus, while examining existing folders to categorize an information item, and selecting a folder, affiliation had an impact on the process.

Appropriateness. During the process of organizing personal information, appropriateness or inappropriateness of categorization affected the decision to keep categorization or make changes in the (6) Categorization stage. On the one hand, when the participants thought the folder where they placed an information item was an appropriate place to save that information item, they kept the categorization. For example, when the researcher asked the participant in the second interview the reason why he/she did not make any changes to the information item, the participant said:

“It seems like the proper place for it.” (Pre-qual Graduate 2)

In addition, participants used the expression ‘happy’ or ‘satisfied’ to describe the appropriateness of categorization when the researcher asked the reason for keeping categorization. For instance, one of the participants said:

“I’m happy with this organization.” (Pre-tenured Professor 1)

On the other hand, when participants thought the categorization was inappropriate, they decided to re-categorize information items. For example, one of the participants said:

“I stored it wrong, and I better move it at some point.” (Tenured Professor 1)

Availability. Although there were only few occurrences of this code, sometimes organizational decisions were made depending on the availability or unavailability of a folder. In the (3) Temporary Categorization stage, sometimes information items were saved into a temporary location simply because the participant did not have any folder for that information item. For instance, when the researcher asked the participant why he/she decided to save an information item into a temporary location, the participant said:

“I didn’t have a folder yet for that company.” (Late Undergraduate 3)

On the other hand, in the (5) Selection/Modification/Creation stage, sometimes participants saved an information item into a certain folder because it was the only category he/she had available. For example, when the researcher asked why the participant saved an information item into a certain category, the participant said:

“That’s because I don’t really have anywhere else.” (Early Undergraduate 1)

Format. The physical characteristics of an information item were one of the primary factors that influenced the process of organizing personal information. More specifically, the code occurrences for ‘format’ of an information item showed the seventh highest number of occurrences among 19 factors. The analysis of the results showed that this factor affected the (2) Identification stage, (4) Examination/Comparison stage, and (5) Selection/Modification/Creation stage. Especially, the ‘format’ factor heavily impacted on the identification of an information item so that among 41 occurrences of the ‘format’ code, 26 of them (63.4%) influenced the (2) Identification stage. In the (2) Identification stage, participants often typified information items by focusing on the digital format of an information item. Specific examples are as follows:

“They were photos.” (Tenured Professor 4)

“It’s a PDF file.” (Pre-qual Graduate 3)

“It’s a Word file.” (Tenured Professor 3)

As shown in these examples, while there were various aspects of an information item, sometimes participants simply identified an information item based on its digital format (“photos”, “PDF file”, or Word file”).

Format of an information item also influenced the (4) Examination/Comparison and (5) Selection/Modification/Creation stages of the personal information organization

process. For example, while examining and reviewing existing folders, one of the participants said:

“There’s a ‘My Photos’ icon that you can save all your pictures in.” (Late Undergraduate 3)

Here, in this example, format of an information item (“pictures”) influenced the examination and review of an existing folder (“My Photos”).

Similarly, in the (5) Selection/Modification/Creation stage, when the researcher asked why the participant decided to select a certain category to save an information item, the participant said:

“Because it’s a picture. It’s a ‘Picture’ folder. It’s a folder for pictures.” (Pre-qual Graduate 3)

Another example is as follows:

“That goes under ‘Teaching’ and then the ‘Class’ and then in ‘Slides’... I was trying to think whether I should call it notes or slides. But I usually think of notes as being Word documents.” (Post-qual Graduate 2)

In this example, the ‘format’ of an information item affected both (4) Examination/Comparison and (5) Selection/Modification/Creation stage.

Messiness. In Section 5.2.3, it was explained that participants cognitively felt that it was messy when information items were not organized. The analysis of the data showed that this mental ‘messiness’ affected both the (3) Temporary Categorization stage and the (6) Categorization stage, so that participants decided to either re-categorize an information item or delete it when they felt it was messy. For instance, in the (3)

Temporary Categorization stage, while explaining about the possibility of re-categorizing an information item that was saved in the temporary location, one of the participants said:

“Just because I try and I want to keep my ‘Desktop’ screen clean so if there’s files I’m not accessing as much, I usually move them to a folder.” (Late Undergraduate 3)

Here, the participant mentioned that he/she re-categorized an information item to eliminate mess in the temporary location (“keep my ‘Desktop’ screen clean”). In addition, while describing the possibility of deleting an information item that he/she saved on the temporary location, one of the participants said:

“I don’t want to clutter too much.” (Pre-tenured Professor 2)

In this example, mental messiness affected deleting decision in the (3) Temporary Categorization stage.

Similarly, in the (6) Categorization stage, ‘messiness’ affected the decision of both re-categorization and deletion of categorized information items. For instance, participants said:

“I did it partly because there was too much mess on my screen.” (Tenured Professor 1)

“I only needed it for homework and when I was done with it I didn’t need it crowding up my space.” (Late Undergraduate 1)

Here, in the first example, the participant re-categorized an information item that was placed into a category by splitting this category into three subfolders. When the researcher asked the reason why he/she did it, the participant explained that ‘messiness’ was one of the factors that affected this re-categorization decision (“too much mess on

my screen”). In the case of the second example, the answer from the participant showed that eliminating the messiness was one of the reasons why he/she deleted a file that was placed into a category (“I didn’t need it crowding up my space”).

Necessity of differentiation. Sometimes the necessity of differentiating certain information items from other information items influenced the process of organizing personal information. In particular, this factor led participants to re-categorize an information item in the (6) Categorization stage that resulted in modification of existing categories in the (5) Selection/Modification/Creation stage. One way of differentiating certain information items from other information items was by creating a new category. For instance, participants said:

“I made the new ‘Stats’ one because I wanted to make sure I had the stats separate.” (Early Undergraduate 3)

“It’s a totally new job, so I don’t want to conflict with my existing folders.” (Pre-qual Graduate 1)

In both cases, participants created a new category to keep certain information items separate from other information items. Another way of differentiating certain information items from others was by creating subfolders within a folder. Following are the examples:

“I put it in ‘Exam 2’ folder just so, I can just distinguish between what I need to know for, what I need to know for Exam 3.” (Late Undergraduate 2)

“Because in ‘University 1’ (folder) there are other stuff like class or administrative documents and everything. So I need to distinguish it from other academic things or administrative things.” (Pre-qual Graduate 3)

In the first example, the participant created ‘Exam 1’, ‘Exam 2’, and ‘Exam 3’ subfolders within a ‘Course 1’ folder so that he/she could differentiate course materials for each exam. In the second example, the participant created a ‘Works’ folder within the ‘University 1’ folder to differentiate information items that are related to works from other information items, such as those for classes or administrative information items.

Number of files. The number of files also influenced the process of organizing personal information. In particular, ‘anticipation of more files’, ‘few number of related files’, and ‘many number of related files’ impacted the process of personal information organization. In the case of ‘anticipation of more files’, it affected the decision to create a new folder in the (5) Selection/Modification/Creation stage. Participants decided to create a new category when they anticipated having more related files. For example, one of the participants said:

“If you are going to get a lot of similar things, and then you may create a folder.”

(Pre-tenured Professor 2)

Also, when the researcher asked why he/she created a new category, another participant responded:

“Just because I figured there would be more things that would be going into that folder.” (Late Undergraduate 3)

In the case of ‘few number of related files’, participants decided not to create a new category or subcategory within a category when they had only a few related information items. For example, participants said:

“It was, I think, only two pictures, so I didn’t need a whole folder for it.” (Late Undergraduate 3)

“But then I will create folders for each one. I just don’t have that many files yet, so I didn’t need to.” (Tenured Professor 3)

“So eventually when I have sufficient numbers of articles related to this. I mean I have files of references, I will actually move it into a separate folder called ‘Course 1’, but I’m not ready to do that yet.” (Tenured Professor 2)

In contrast, when participants have ‘many number of related files’, they decided to either create a new category or modify an existing category by creating subcategories. For example, participants said:

“Because it was an event that had more than five pictures, so I knew I had to create a folder to keep them all together.” (Late Undergraduate 1)

“I had like, like more than five or six versions of them, so I decided to create just a folder to organize all the related things.” (Post-qual Graduate 1)

This finding corresponds with Ravasio, Schär, and Krueger’s (2004) study, which reported that participants create new folders when there were several files on the same subject (p. 164). Barreau (1995) also mentioned a participant who kept creating subfolders as the amount of information grew (p. 337).

In a similar vein, when participants have ‘many number of related folders’, they created a superordinate category to group similar categories. The following is an example:

“Because I realized that there many different traveling pictures folders lying around in my ‘Picture’ folder so I just wanted to collect them together in one folder.” (Pre-qual Graduate 3)

Thus, as shown above, ‘number of files’ influenced the process of organizing personal information, especially on the (5) Selection/Modification/Creation and (6) Categorization stages. In addition, this factor was directly related to ‘accessibility’ and ‘necessity of differentiation’ factors. When there were only a small number of information items, it was easy to access them so that participants did not need to group them together, or make distinctions among information items in a category. However, when the number of files grew, or when participants expected to have more files in the future, it became difficult to access information items, so that they needed to group them together and differentiate some of them from other information items to enhance accessibility. For instance, participants said:

“There were too many of them and it starts (to get) hard to see things because I got intertwined, hard to search and find things.” (Tenured Professor 1)

“Just because I figured there would be more things that would be going into that folder so that it’s easier if I need to have, say, a ‘Writing’ (folder) for it, I could just access them all in the same place.” (Late Undergraduate 3)

Purpose. The purpose of an information item was obviously the most influential factor for process of organizing personal information; among 19 factors, ‘purpose’ showed the highest number of occurrences. The analysis of the data showed that the ‘purpose’ factor affected four different stages of the personal organization process: (2) Identification; (4) Examination/Comparison; (5) Selection/Modification/Creation; and (6) Categorization stages.

In particular, this factor heavily influenced the (2) Identification stage; among 250 code occurrences of the ‘purpose’ factor, 104 of them affected the (2) Identification stage

(41.6%). In this stage, participants frequently typified an information item based on the main purpose of that information item. For example, participants said:

“It’s for class.” (Early Undergraduate 1)

“That was for my Exam 2” (Late Undergraduate 2)

“This is for the class I’m teaching tonight.” (Post-qual Graduate 2)

“This is for the new conference.” (Pre-tenured Professor 2)

As shown above, often participants focused on the main purpose of the information item and typified it.

In addition, the ‘purpose’ of an information item influenced the (4) Examination/Comparison stage. In particular, the purpose of an information item affected this stage when participants reviewed existing folders. For example, participants said:

“I have individual folders for each of my classes.” (Late Undergraduate 1)

“Everything related to my dissertation is here.” (Post-qual Graduate 2)

“So I have another folder, that’s just for the research picture.” (Pre-tenured Professor 2)

As shown above, participants often grouped information items together based on the purpose of an information item so that when they reviewed existing categories to find relevant folders to categorize an unorganized information item, the ‘purpose’ of an information item influenced the process. The purpose of an information item also influenced the (4) Examination/Comparison stage when participants assessed similarities and differences between new and existing files. For instance, participants said:

“So it’s a conference thing, but it’s not really a publication or a paper. As I say, it’s a tutorial. So I saved it under teaching.” (Pre-tenured professor 1)

“I could put it under ‘Reviews’ but, it’s really not quite reviews, it’s the assignments of what I’m supposed to read, and what everybody else is supposed to read.” (Tenured Professor 1)

Here, while assessing similarities and differences between new and existing information items to decide which category to organize an information item into, the purpose of both new and existing information items affected the process.

The purpose of an information item also influenced the (5) Selection/Modification/Creation stage. Among 250 code occurrences of the ‘purpose’ factor, 88 impacted this fifth stage of the process (35.2%). Examples are as follows:

“They’re all related to the same class, so I figured I should have them all in the same folder.” (Late Undergraduate 3)

“It’s for my dissertation. So it goes in the ‘Dissertation’ folder.” (Post-qual Graduate 2)

“Anything that had to do with that TA-ship or whatever you would call it, I put into that folder.” (Pre-qual Graduate 2)

As shown in these examples, participants often grouped information items together and separated certain information items from other items based on the main purpose of an information item. In other words, when categorizing an unorganized information item, participants often selected a folder which contained information items for the same purpose as the unorganized information item. Thus, when assessing similarities and differences between new and existing files, when the purposes of those information items were the same, this purpose was exaggerated as a similarity, so that the items could be

grouped together. However, when they were for different purposes, the purposes of these information items were inflated as differences, so they were separated from each other.

In addition, in the (5) Selection/Modification/Creation stage, often participants created a new category when they had information items for a new purpose, so that they could save information items that were related to the new task in the same category. In this case, participants rarely could find an existing category that contained information items for the same purpose, which made them create a new folder. For example, participants said:

“Because I didn’t have one specifically dedicated to this particular committee and I needed one.” (Tenured Professor 2)

“Because these were the first, I guess, documents that I was saving to my laptop for the summer, so trying to be organized so I figured I’d make a new folder for summer session.” (Late Undergraduate 3)

“[It is] Different from my Conference 1 poster, so that’s why I created a new folder called ‘Project 1’ as it’s the new job task.” (Pre-qual Graduate 1)

“This is the conference we start to work on right now.” (Pre-tenured Professor 2)

In all of the examples above, participants created a new folder because they got a new task such as a new committee work, summer session, new project, and new work for a conference.

Not surprisingly, in the (6) Categorization stage, participants often decided to keep the categorization when the purpose of a categorized information item was the same. Thus, when the researcher asked participants why they decided to keep the categorization, one of the participants said:

“Because that’s for that conference paper, that project.” (Pre-tenured Professor 2)

Similarly, when the researcher asked if there was a possibility of making changes to the categorization, one of the participants said:

“No, not likely, no. It’s very related to this task.” (Tenured Professor 1)

Thus, participants decided to keep the categorization when the purpose of an information item had not been changed.

However, when the purpose of an information item changed, participants made changes to the categorization or decided to make changes in the future. For example, participants said:

“When I change my dissertation into articles or a book, they’ll all go under there also.... because then they become writings...Once I try to publish them, they’ll be under Writings.” (Post-qual Graduate 2)

“I will actually touch it next time when I teach this class because I’ll be looking for what was that ‘Quiz 1’ last time, so I’ll look at that, and actually first thing I’ll do is just copy that file to the new class, and then open it, do some editing and that’ll be the quiz for class next time I teach it.” (Pre-tenured Professor 1)

In the first example, the participant said that when he/she published an article or a book from his/her dissertation, he/she would place that information item into the ‘Writings’ folder instead of the ‘Dissertation’ folder. In this case, the change of the purpose of an information item was one of the factors that influenced the decision. In the second example, the participants said that in the future, when he/she teaches the same class he/she was teaching at the time of interviewing, he/she would copy an information item and move it to the new class folder. In this case, although the information item will

remain in the folder in which the participant originally placed it, it also will be copied and moved into another folder. In other words, the participant decided to make a change to the organization in the future when the purpose of the information item changes. Thus, changes in the purpose of information items influenced the (6) Categorization stage of the process of organizing personal information.

Related person. A person who is related to an information item was also one of the primary factors that influenced the process of organizing personal information. More specifically, the code occurrences for ‘related person’ showed the sixth highest number of occurrences among 19 factors.

In particular, this factor heavily influenced the (2) Identification stage; with 24 of the 42 occurrences of the ‘related person’ code (57.1%). In this stage, participants often identified an information item in relation to someone. This included (1) when they worked on an information item with someone, (2) when it was someone’s work, (3) when the contents of an information item were related to someone, and (4) when an information item was for someone or (5) from someone. Specific examples are as follows:

“My doctoral student is working with me. So we are trying to come up with a conference paper.” (Pre-tenured Professor 2)

“This is the most recent version of his thesis.” (Tenured Professor 1)

“They were photos of my son's graduation.” (Tenured Professor 4)

“It’s like an exercise for students to do.” (Pre-qual Graduate 2)

“These are all the task my advisor gave to me.” (Post-qual Graduate 1)

In the first example, a person who worked on an information item together with the participant (“My doctoral student is working with me”) influenced identifying an information item. In the second example, the participant identified an information item by recognizing and stating that it was someone’s work (“his thesis”). In the third example, the topic of an information item was related to someone (“my son’s graduation”), while it was for someone in the fourth example (“for students”). In the last example, an information item was from someone (“my advisor gave to me”). Thus, the ‘related person’ factor influenced the identification of an information item in various ways.

In the (4) Examination/Comparison stage, ‘related person’ affected the process as participants reviewed existing folders which grouped information items based on the name of a person who was related to those information items. Thus, when participants reviewed the existing categories or when they compared new and existing information items to find relevant folders to categorize an unorganized information item, ‘related person’ influenced the process. For example, while reviewing existing folders, participants said:

“I have one called ‘Daughter’s name’ which is anything related to my daughter.”

(Tenured Professor 3)

“If I have a graduate student who is particularly productive, I would have a folder named after him or her.” (Tenured Professor 2)

This factor also influenced the (5) Selection/Modification/Creation stage. When the researcher asked why a participant selected one of the existing folders, he/she said:

“There is folder for graduate students, graduate students I’ve had. And so he would be inside of that folder, his name would be there.” (Tenured Professor 1)

Participants sometimes created a folder by the name of a person who is related to the information items. For instance, one of the participants said:

“He just sends it to you without any email body message. Just an attachment and nothing. And in this way, I realized that if I want to go back to whatever he sent to me, it’s very hard. So I create a folder for him. Every time, I download his documents into that folder, um, I mean for this way, I can do a better search.”

(Pre-tenured Professor 2)

Here, the participant created a new category to group information items that were sent from a particular person (“I create a folder for him”). Thus, as shown in these examples, participants often grouped information items together and separated certain information items from other items based on the name of a person who is related to an information item.

Reminder. Interestingly, ‘reminder’ was one of the factors which affected the process of organizing personal information. More specifically, sometimes participants decided to save an information item in the temporary location with or without a category so that they could be reminded of certain tasks. Because the temporary location is usually a place which participants can either easily see or frequently access, placing an information item into these locations worked as a visual reminder for the participants.

In the (3) Temporary Categorization stage, participants saved an information item into a temporary location without a folder to remind them of certain tasks they needed to do. For instance, participants said:

“It’s still on the ‘Desktop’ because it says ‘Naughty, naughty, naughty, Tenured Professor 1, it’s still here, you should read this.’ And notice it’s not in the folder

or anything. It's really standing out there and says 'You better do something about this'. And that's one of my procedures. Those things that are sitting here like this, they're just files that say 'Do something'." (Tenured Professor 1)

"[I] put on my 'Desktop' so that I can remind myself that I have to do it." (Late Undergraduate 1)

In the (5) Selection/Modification/Creation stage, participants also created a new folder in a temporary location or selected one of the existing folders in the temporary location to be reminded of certain tasks. For example, participants said:

"I put it on, in a folder on my 'Desktop' to remind me I have papers to give him feedback on it." (Tenured Professor 1)

"I would check this folder with some degree of frequency and be reminded of it whereas I would have to make a more deliberate of acts to go to 'Conferences' (folder) to search the organization, find, basically what I'm trying to do is minimize the number, then depth which I have to click in order to find what I need for the moment." (Tenured Professor 2)

"I'm no longer teaching that course this semester. And I thought if I were to put it in the 'Course' folder, I might – in a sense it's not active in my mind, I might overlook it." (Tenured Professor 4)

Here, in the first two examples, participants mentioned the reason why they selected a folder in a temporary location, i.e., to be reminded of certain tasks, while the participant in the last example explained that if he/she saved an information item into one of the existing categories in the folder structures, he/she would not be reminded of the task he/she

needed to do. Thus, ‘reminder’ was one of the factors that impacted on the process of personal information organization.

Source. Where the participant obtained an information item also influenced the process of organizing personal information; the code occurrences for the ‘source’ of an information item showed the tenth highest number of occurrences among 19 factors. However, this factor influenced only one stage of the process, which is the (2) Identification stage. When identifying an information item, participants often typified an information item while focusing on the source of the information item. For example, participants said:

“Stuffs I save from the internet.” (Early Undergraduate 1)

“I got the file from Google Scholar.” (Pre-qual Graduate 1)

As shown in these examples, where the participants obtained an information item affected the (2) identification stage of the personal information organization process.

System. The automated system or the default settings of certain software, web browsers, or personal devices inevitably influenced the process of organization, so that regardless of the intentions of the participants, information items were saved or organized in certain ways because the automated system of certain software, web browser, or device required it.

In the (1) Initiation stage, sometimes software on a device automatically saved an information item into the personal device when the participants connected the device with his/her personal computing device. For example, participants said:

“I plug it and it automatically uploads anything new.” (Early Undergraduate 3)

“Yes, that’s default. I don’t do anything with that. So Amazon automatically puts everything into ‘My music’ (folder).” (Post-qual Graduate 2)

Similarly, in the (3) Temporary Categorization stage, certain personal devices or web browsers automatically saved certain information items into a temporary location.

For instance, participants said:

“When you download something, it automatically puts it in ‘Download’ folder.”

(Pre-tenured Professor 2)

“When I download something, unless I change it to like my hard drive or my

‘Desktop’ or something, it’s gonna automatically save in my ‘Downloads’

folder.” (Late Undergraduate 1)

In both examples, any information items which participants downloaded were automatically saved into the ‘Downloads’ folder, which was used as a temporary location to save an information item.

In a similar vein, some systems automatically created a folder, or organized information items into folders so that it affected the (5) Selection/Modification/Creation stage. For example, participants said:

“Mac automatically created a folder on my ‘Desktop’.” (Pre-tenured Professor 2)

“It automatically makes the folder with the artist’s name on it.” (Post-qual

Graduate 2)

Thus, system was one of the factors that influenced the process of organizing personal information.

Time. The time dimension of an information item such as year, season, semester, week, or day was one of the prominent factors that affected the personal information

organization process. This factor influenced the (2) Identification, the (4) Examination/Comparison, and the (5) Selection/Modification/Creation stages, and it showed the fourth highest number of occurrences among 19 factors. The analysis of the data shows that the code occurrences are distributed almost equally among three stages. Among 66 code occurrences, 25 (37.9%) belonged to the (2) Identification stage, 18 (27.3%) belonged to the (4) Examination/Comparison stage, and 23 (34.8%) belonged to the (5) Selection/Modification/Creation stage.

In the case of the (2) Identification stage, the time dimension of an information item affected the way participants identified an information item. For example, participants said:

“This is the paper that I worked on last year, last summer.” (Post-qual Graduate 1)

“It’s my reading list for summer 2012. It’s reading list for the summer vacation.”

(Pre-qual Graduate 3)

“It’s the students in my class for the fall.” (Tenured Professor 3)

As shown in these examples, the time dimensions of an information item were used in identifying an information item. However, unlike many other factors, in the case of the ‘time’ factor, it was often not used independently but along with other factors. In the first two examples, both the ‘purpose’ of an information item (“the paper that I worked on” and “reading list”), and the ‘time’ dimension of an information item (“last year, last summer” and “summer 2012”), were used together in identifying an information item. Likewise, in the last example, both ‘related person’ of an information item (“the students in my class”) and the ‘time’ dimension of an information item (“for the fall”) were used in identifying an information item.

In the (4) Examination/Comparison stage, the time dimension of an information item impacted on this stage when participants reviewed existing folders. For example, participants said:

“I have like categories based on like seasons, so I have like, from high school, I have ‘Senior’, ‘Junior’, and then from College I have either by a semester, so like ‘Sophomore spring’, or I have ‘Christmas break’, ‘Summer 11’, like I have all the different seasons basically.” (Early Undergraduate 2)

“I keep all of my spring semester classes in one place just to make finding things easier. So, I have subfolders in the ‘Spring semester’ folder.” (Late Undergraduate 1)

“The subject name is ‘Class 1’, and then under that, things are organized according to semester in which I teach.” (Pre-tenured Professor 1)

As shown in these examples, participants often grouped information items together based on the time dimension of an information item (such as year, semester, or season) so that when they reviewed existing categories to find relevant folders to categorize an unorganized information item, the time dimension of an information item influenced the process.

The time dimension of an information item also influenced the (5) Selection/Modification/Creation stage of the process. When the researcher asked why the participant selected a certain category to organize an information item, participants said:

“Just because it’s related to my summer class with the rest of my summer stuff.”
(Late Undergraduate 3)

“I saved it there because it was taken on the same day as the other pictures, and it has to do with them.” (Late Undergraduate 1)

As shown in these examples, the time that is associated with an information item often affected the way participants organized their personal information.

Time availability. Whether participants had enough time to make appropriate organizational decisions or not also affected the process of organizing personal information. It is important to note that this factor is different from the ‘time’ factor explained previously. The ‘time’ factor indicates the time dimension associated with an information item such as year, semester, season, week or day. The ‘time availability’ factor refers to the length of time that was available to the participants when they were organizing personal information. This factor influenced the (3) Temporary Categorization, the (5) Selection/Modification/Creation, and the (6) Categorization stages. This factor heavily influenced the (3) Temporary Categorization stage, representing 11 out of the 19 occurrences of ‘time availability’ (57.9%).

In the (3) Temporary Categorization stage, participants often decided to save an information item into a temporary location when they did not have enough time to make an organizational decision. For example, participants said:

“I probably will [create a folder] on Monday, but I just was in a rush when I was working on it, so I just saved it quickly to my ‘Desktop’.” (Late Undergraduate 3)

“A lot of it is just keeping things accessible and relatively easy to find without taking the time to put them away because that time is – I just don’t have that.”

(Tenured Professor 3)

In the (5) Selection/Modification/Creation stage, sometimes participants placed an information item in an inappropriate folder because they did not have enough time to make an appropriate decision. For example, when the researcher asked why a participant placed an information item into a less relevant category, he/she said:

“Because I was in a hurry.” (Tenured Professor 1)

When the researcher asked what he/she would do if not in a hurry, the participant said:

“I would’ve moved it there immediately.” (Tenured Professor 1)

Similarly, in the (6) Categorization stage, participants postponed a re-categorization decision when they did not have enough time to re-categorize information items. For example, participants said:

“Hopefully I will find more time next week. I am going to take care of all these files and put them in the right folder.” (Post-qual Graduate 3)

“And then I will probably create three subfolders under that... It’s just that I haven’t done that yet because I’ve been too busy to bother doing that.” (Tenured Professor 3)

As shown above, participants often delayed re-categorization because they lacked time. Thus, ‘time availability’ affected the process of organizing personal information.

Topic. What an information item is ‘about’ was one of the primary factors that influenced the process of organizing personal information. The code occurrences for the ‘topic’ of an information item represented the fifth highest number of occurrences among 19 factors.

In the (2) Identification stage, participants often identified an information item by the subject of an information item or what was contained in that information item.

Participants said:

“So I downloaded this which is an article about peer friendships.” (Tenured Professor 2)

“It’s related to my area. So I do book history. Well, I do book culture and reading practices. It’s related to that.” (Post-qual Graduate 2)

“It’s about Facebook use and civic and political engagement.” (Pre-qual Graduate 1)

As shown in these examples, what an information item is about (i.e., topic of an information item) influenced identifying information items to organize them.

In the (4) Examination/Comparison stage, although there were few occurrences, the analysis of the data showed that participants sometimes grouped information items based on the ‘topic’ of an information item so that when they reviewed existing folders, the ‘topic’ of an information item influenced the process. For example, while reviewing existing folders, one of the participants said:

“Information technology is something that I have an interest in, so I have a folder in there called ‘IT’, so papers that I write in relation to IT get stored there.”
(Tenured Professor 2)

In the case of the (5) Selection/Modification/Creation stage, ‘topic’ obviously affected the process as participants decided to select or deselect a certain category because of the ‘topic’ of an information item. For example, when the researcher asked why he/she selected certain category to organize an information item, a participant said:

“[I was] Trying to organize the same topic.” (Pre-qual Graduate 1)

Thus, the ‘topic’ of an information item was a factor that influenced the decision of selecting a certain folder to organize an information item. On the other hand, when the researcher asked another participant why he/she decided to select a certain folder instead of another folder, the participant said:

“Because they are different topics, completely different topics.” (Post-qual Graduate 3)

Thus, the ‘topic’ of an information item also influenced the decision of grouping or separating information items.

This factor also influenced the decision to create a new folder. Participants sometimes created a folder based on the topic. For instance, one of the participants said:

“I named the [folder] ‘Translation democratization’ because translation is the kind of work that I’m doing. And democratization is the subject of the work.” (Pre-qual Graduate 3)

Type. The general type or the genre of an information item also influenced the process of organizing personal information. This factor affected the (2) Identification and the (4) Examination/Comparison stages.

In particular, the ‘type’ of an information item affected the (2) Identification stage, representing 28 among the 34 occurrences of ‘type’ (82.4%). In this stage, participants identified an information item based on its ‘type’. Specific examples are as follows:

“So it's an article.” (Tenured Professor 4)

“That’s another set of minutes.” (Pre-qual Graduate 2)

“This is the actual template.” (Post-qual Graduate 3)

As shown above, the ‘type’ of an information item (article, minutes, or templates) affected the identification of an information item.

In the (4) Examination/Comparison stage, the ‘type’ of an information item affected the process as participants reviewed existing folders which grouped the same type of information items. Thus, when participants reviewed the existing categories or when they compared new and existing information items to find relevant folders to categorize an unorganized information item into, the ‘type’ of an information item influenced the process. For example, while reviewing existing folders, participants said:

“I have a folder called ‘Articles’.” (Tenured Professor 2)

“This one, No. 10. I had to think about where to put that. Because I do have another thing called ‘Articles’.” (Post-qual Graduate 2)

Use. The ‘use’ factor was one of the primary factors that affected the process of organizing personal information. This factor had the second highest number of occurrences among 19 factors. Additionally, this factor was the only factor that influenced all six stages of the process. However, it is important to note that the ‘use’ factor is different from the ‘purpose’ factor. More specifically, the ‘purpose’ factor indicates what an information item is ‘for’ (such as for dissertation, for a class, or for an exam). The ‘use’ factor indicates whether participants are currently using, not using, will be using, or will not be using an information item, regardless of the purpose of an information item.

The ‘use’ factor affected the (1) Initiation stage, when participants decided to save or obtain an information item to use it. The analysis of the data showed that participants

often saved information items when they needed to use information items. For example, participants said:

“It’s an assignment that I had to hand in. So that’s why I had to save it.” (Late Undergraduate 2)

“I downloaded this so that I could study.” (Late Undergraduate 1)

“I just had to print it and send it to the Visa agent.” (Pre-tenured Professor 1)

In addition, participants saved information items when they needed to use information items in the future, or if there was a possibility of using those information items in the future. For instance, participants said:

“You asked why did I save it? For future use. All of these are for future use.”
(Tenured Professor 3)

“I just was saving it for reference purposes in case I needed to look at it again.”
(Pre-qual Graduate 2)

“It’s something that maybe someday, somewhere I’ll use in some presentation. But I don’t have any immediate use or need for it, so I just saved it.” (Pre-tenured Professor 1)

In the (2) Identification stage, the ‘use’ factor influenced the way participants identified an information item. For example, participants said:

“These were really instant files.” (Post-qual Graduate 1)

“That’s just for one term use.” (Pre-tenured Professor 2)

In these examples, participants identified information items based the ‘use’ of an information item. In the first example, the information items were identified as files to be used for a short period of time (“instant files”), while an information item was identified

as a file to be used once (“one term use”). Thus, the ‘use’ factor influenced the (2) Identification stage.

In the (3) Temporary Categorization stage, participants decided to save an information item into a temporary location because the information item was for short term use. For example, participants said:

“If I do not expect that I would use it again, I would put it in ‘Desktop’.” (Pre-tenured Professor 2)

“Those are really instant things, so I cannot think of any folders to save them.” (Post-qual Graduate 1)

“The reason I didn’t save it in any other folder, except for my ‘Downloads’ is because I will never need it again.” (Late Undergraduate 1)

In addition, participants saved an information item into a temporary location when there was an ongoing use of an information item. Examples are as follows:

“Because I’m working on things at the moment, I’m keeping things on the ‘Desktop’.” (Tenured Professor 2)

“I’ll put it on my ‘Desktop’ if you know, um, constantly using it?” (Early Undergraduate 1)

“I didn’t put it back because it was a current file that I needed, so I just saved it to ‘Desktop’ for the time being.” (Tenured Professor 3)

However, when there was no future use for an information item, participants deleted it. When the researcher asked participants why they deleted information items that were saved in the temporary locations, participants said:

“I didn’t think I needed it anymore.” (Tenured Professor 1)

“I don’t need it anymore...It’s not something I see myself doing in the future, so.”

(Late Undergraduate 1)

“I didn’t need it anymore.” (Post-qual Graduate 1)

Hence, the ‘use’ factor impacted on the (3) Temporary Categorization stage in various ways.

Although there were only few cases in the (4) Examination/Comparison stage,, ‘use’ worked as a factor, especially while reviewing and comparing relevant categories to decide where to save an information item. For example, participants said:

“It’s not school information, but I probably will, might go look at it later, so I wouldn’t just save it into ‘Documents’, either.” (Early Undergraduate 2)

“Because the only problem with putting it in my G-mail is then it would feel like something I need to do right away. And that’s why I’ve saved it there so that I don’t have to see it when I open my e-mail – my inbox.” (Post-qual Graduate 2)

In the (5) Selection/Modification/Creation stage, one of the reasons why participants decided to select a category was similar to what occurred in the (3) Temporary Categorization stage. Participants decided to select a category in the temporary location when they have an ongoing use of an information item. For example, participants said:

“It’s on my ‘Desktop’ because I’m working on it. It’s, I have, I create folders on my ‘Desktop’ for things that I’m working on.” (Tenured Professor 1)

“But again, because I’m working on things at the moment, I’m keeping things [in folders] on the ‘Desktop’.” (Tenured Professor 2)

Thus, although files were saved without a folder in the (3) Temporary Categorization stage, and they were saved within a folder in the (5) Selection/Modification/Creation stage, the way the ‘use’ factor influenced both these stages was almost the same.

In this stage, the ‘use’ factor also affected the modification of the categorization. When there was no ongoing use of an information item, participants decided to make modifications to the categorization either by creating a subordinate category or superordinate category. For example, participants said:

“After each quiz, I don’t need that information so I just store it separately.” (Late Undergraduate 2)

“I actually at the end of the semester will just combine all my folders because I don’t need them for classes anymore.” (Late Undergraduate 1)

“I may just create a new folder, maybe just creating the bigger folder called ‘2012 spring’ or summer or whatever and then just dump those files into that new folder... because the old job is done so I possibly will not open the folder often than I used to do that.” (Pre-qual Graduate 1)

In particular, the ‘use’ factor heavily influenced the (6) Categorization stage, in that it represented 72 among the 175 occurrences of the ‘use’ factor (41.1%), affecting decisions in various ways.

First, when there was ongoing use of an information item, participants decided to keep the categorization. Thus, when the researcher asked participants in the second interview why they did not make any change to the categorization, participants said:

“It's an ongoing project.” (Tenured Professor 4)

“I’m still on the committee and meeting through the summer so I haven’t moved the folder to anywhere else yet.” (Tenured Professor 2)

“It’ll probably stay there until we go on the vacation just because I’ll be adding more stuff to it.” (Late Undergraduate 3)

In addition, when the participants did not use an information item, they also kept the categorization. When the researcher asked participants in the second interview why they decided to keep the categorization of certain information items, participants said:

“It’s still there. I haven’t worked on that, yet.” (Pre-tenured Professor 2)

“It’s the other student who needs to now edit it. So then, I won’t move it, too, because I’m not doing any editing.” (Pre-tenured Professor 1)

“I just wanted to store them there and then I didn’t change it.” (Pre-qual Graduate 3)

Thus, participants kept the categorization when there was no use of the file.

When there was no ongoing use of an information item, participants often decided to re-categorize an information item.

“I just don’t need for the next quiz. So I just put it away in a subfolder.” (Late Undergraduate 2)

“So once I don’t use those files or papers, then I try to create their new folder called ‘2012 spring’ for example, I try to move those things into that folder.” (Pre-qual Graduate 1)

“Once I’m finished with it and I don’t need it immediate, more present for myself, I will put it in the deeper folder.” (Tenured Professor 2)

However, when there is a possible future use, participants sometimes decided to move it to another device for archiving. For instance, participants said:

“When I do actually want to save something permanently there is another system I use.” (Early Undergraduate 3)

“So I have a backup and, yeah, certainly on my laptop I need to move it over there.” (Tenured Professor 4)

Thus, possible future use of an information item influenced the participants’ decisions of whether to move it to a more permanent location or not in the (6) Categorization stage.

Not surprisingly, when there was no future use of an information item, participants often deleted the file. Thus, when the researcher asked why they decided to delete an information item that was placed into a category, participants said:

“I only needed it for homework and when I was done with it I didn’t need it crowding up my space.” (Late Undergraduate 1)

“Because I found out that I didn’t need it for work anymore, so I just put it in my recycle bin.” (Late Undergraduate 3)

“I just felt like, I wasn’t ever really like going back and looking at them, it’s just like, stupid, so I’m just gonna erase it.” (Early Undergraduate 2)

As shown above, the ‘use’ factor influenced on the process of organizing personal information from the (1) Initiation stage to the (6) Categorization stage.

Value. The value of an information item, which refers to the personal meaning of an information item to the participants, also influenced the personal information organization process. In addition, although there were just few occurrences in each stage, this factor affected five different stages of the process: the (1) Initiation stage; the (2)

Identification stage; the (3) Temporary Categorization stage; the (5) Selection/Modification/Creation stage; and the (6) Categorization stage.

In the (1) Initiation stage, participants sometimes saved information items because those information items were interesting or funny. For example, participants said:

“It was funny so I wanted to save it to my own cameral roll, and so I did that.”

(Early Undergraduate 2)

“Any of the articles that I found really interesting I would basically open, bookmark and then I save them.” (Early Undergraduate 3)

In the (2) Identification stage, some information items were identified based on the value of those information items. For instance, participants said:

“That was a picture probably of something like, sappy or like mushy.” (Early Undergraduate 2)

“So it is one of my favorite TV shows.” (Pre-qual Graduate 1)

In the first example, the participant identified an information item based on the format (“picture”) and the value (“sappy” and “mushy”) of an information item. Similarly, in the second example, the participant identified an information item based on the type (“TV shows”) and the value (“my favorite”) of an information item.

In the (3) Temporary Categorization stage, one of the reasons why participants decide to save an information item into a temporary location instead of one of the categories in their folder structures was when an information item was regarded as not very interesting, important, or worthy of saving according to the participants. For example, participants said:

“That’s the main reason for me to just leave those papers into my ‘Downloads’ folder. Those papers are less interesting or less important for (to be organized into) my existing folders.” (Pre-qual Graduate 1)

“I may need to use them in the near future, but they are not that important to save in my, like permanent folder on the internet.” (Post-qual Graduate 1)

“If something were worthy of saving, then I would move it to appropriate places.” (Pre-tenured Professor 1)

In the (5) Selection/Modification/Creation stage, the ‘value’ of an information item influenced the participant’s decisions on selecting a category to organize an information item. More specifically, when the researcher asked participants why they decided to save an information item into a certain category, participants said:

“I save all my photos in iPhoto unless they aren’t important to me.” (Late Undergraduate 1)

“I would consider it like school documents more important than the stuff I usually save into just ‘Documents’ or on my ‘Desktop’ or something.” (Early Undergraduate 2)

In these examples, the ‘importance’ of an information item affected the decision of where to save an information item.

In the (6) Categorization stage, often participants decided to delete an information item when it was regarded as not important or interesting. For example, participants said:

“Probably into the, my trashcan [Laughing]. Because it’s just anything. So I probably will just delete it all.” (Early Undergraduate 2)

“I’ll probably just delete unless one is really interesting.” (Early Undergraduate 3)

To sum up, there were a number of factors that impacted the process of organizing personal information, which were grouped into 19 factors in this study. Often these factors were closely related to each other. For example, when there were a number of information items (“number of files” factor), participants often felt that it was messy (“messiness”), which made them think that they needed to organize their information in certain ways, which included differentiating some of the information items from other information items (“necessity of differentiation”). This process not only eliminates the messiness but also enhances the accessibility of files (“accessibility”). However, each factor had its own unique meaning and ways of influencing the process enough to be separated as an independent factor. Among 19 factors, the most influential factors in terms of frequency were the ‘purpose’, ‘use’, ‘accessibility’, ‘time’, and ‘topic’ factors. The factors that least influenced the personal information organization process were ‘appropriateness’, ‘messiness’, ‘reminder’, ‘availability’, and ‘system’. Nineteen factors listed by their frequency rank are presented in Table 16.

Table 16
Frequency of Occurrences of the Factors Impact on the Process

Rank	Factor	Occurrence
1	Purpose	250
2	Use	175
3	Accessibility	75
4	Time	66
5	Topic	47
6	Related person	42
7	Format	41
8	Type	34
9	Number of files	25
10	Source	24
11	Necessity of differentiation	23
12	Value	20
13	Time availability	19
14	Affiliation	14
15	Appropriateness	12
16	Messiness	10
17	Reminder	9
18	Availability	7
19	System	7

While there were some factors that influenced a single stage of the organization process (10.5%), most factors affected several stages of the process (89.5%), and thus, each stage was influenced by multiple factors. Among 19 factors, the most distributed factors were the ‘use’, ‘accessibility’, ‘value’, and ‘purpose’ factors, which together affected more than half of the stages of the process. The least distributed factors were ‘appropriateness’ and ‘source’ factors, which impacted only one stage of the process. How each factor influenced different stages of the process is presented in Table 17 (an o indicates presence).

Table 17
Ubiquity of Factors on the Process

Factor	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
Accessibility	0		0	0	0	0
Affiliation				0	0	
Appropriateness						0
Availability			0		0	
Format		0		0	0	
Messiness			0			0
Necessity of differentiation					0	0
Number of files					0	0
Purpose		0		0	0	0
Related person		0		0	0	
Reminder			0		0	
Source		0				
System	0		0		0	
Time		0		0	0	
Time availability			0		0	0
Topic		0		0	0	
Type		0		0		
Use	0	0	0	0	0	0
Value	0	0	0		0	0

5.2.6 Influence of Thought Communities

RQ6. How do individuals' thought communities influence the process of organizing personal information?

As described in the Section 3.1, the theoretical framework for this research study is cognitive sociology. Therefore, by using the cognitive sociological perspective, this research study wanted to highlight the influence of thought communities on the process of organizing personal information, which has often been overlooked. As cognitive sociologist Zerubavel (1997) stated:

“Cognitive sociology keeps reminding us that while we certainly think both as individuals and as human beings, what goes on inside our heads is also affected

by the particular thought communities to which we happen to belong. Such communities – churches, professions, political movements, generations, nations – are clearly larger than the individual yet considerably smaller than the entire human race.” (p. 9).

Indeed, the analysis of data showed that academics view reality not only as individuals or human beings but also as members of the academy, which is a particular social environment and thought community.

5.2.6.1 Influence of the Academia Thought Community

In this research study, participants certainly had individual differences as well as some common behaviors in organizing personal information. Both the unique factors and common factors that are not specifically associated with academia affected the process of organizing personal information. However, participants also used their particular socio-mental lenses in viewing reality as academics, which led them to make distinctions between information items in similar ways, which non-members of the academy would not make. Especially, the way participants identified an information item, and the way they assessed and adjusted the mental gaps between information items were heavily influenced by the academia thought community. Therefore, among six different stages of the process of personal information organization, the (2) Identification stage, the (4) Examination/Comparison stage, and the (5) Selection/Modification/Creation stage were parts of the process which were most profoundly influenced by the thought community of participants.

In the case of the (2) Identification stage, among various aspects of an information item, participants focused on certain aspect(s) of it while disregarding other aspects, and typified it as ‘something’ so that they could make organizational decisions. As shown in Section 5.2.5, the aspect(s) participants primarily focused on while identifying an information item included ‘format’, ‘purpose’, ‘related person’, ‘source’, ‘time’, ‘topic’, ‘type’, ‘use’, and ‘value’ of an information item. Table 18 shows the factors that affected the (2) Identification stage, listed in descending order of the frequency of occurrences.

Table 18

Frequency of Occurrences of the Factors Impact on the (2) Identification Stage

Rank	Factor	Occurrence
1	Purpose	104
2	Type	28
3	Format	26
4	Time	25
4	Topic	25
6	Related person	24
7	Source	23
8	Use	15
9	Value	4

Among these factors, some of the factors (such as ‘format’, ‘source’, ‘type’, ‘use’, or ‘value’ of an information item) were not specifically associated with a particular thought community. For instance, regarding ‘format’, participants said:

“That was a picture.” (Early Undergraduate 2)

“It’s a PDF file.” (Pre-qual Graduate 3)

“It’s an Excel file.” (Tenured Professor 3)

Similarly, when the participants focused on the ‘source’ of an information item while identifying an information item, it was not related to a particular thought community. For example:

“Stuff I saved from the Internet.” (Early Undergraduate 2)

“It’s from the same email.” (Pre-qual Graduate 3)

In addition, when the participants focused on the ‘type’ of an information item to identify a file, it was often not particularly associated with a thought community. Examples are as follows:

“It’s a demo.” (Post-qual Graduate 3)

“This is an article.” (Tenured Professor 2)

“It’s a drama.” (Pre-qual Graduate 1)

Also, the ‘use’ of an information item was often not specifically associated with a particular thought community. Examples are as follows:

“I didn’t need it anymore.” (Post-qual Graduate 1)

“That’s what I’m doing right now.” (Tenured Professor 3)

Last, the ‘value’ of an information item also was not particularly related with a specific thought community. For instance, participants said:

“It was just like some cute little picture.” (Early Undergraduate 2)

“Articles that I found really interesting.” (Early Undergraduate 3)

However, when participants focused on the ‘purpose’, ‘time’, ‘topic’, or ‘related person’ of an information item, which showed relatively higher rank in the frequency of occurrences than other factors, it was often related to the academic thought community.

In the case of the ‘purpose’ of an information item, participants often identified an information item based on the primary tasks of academics. For example:

“That one was for that course.” (Early Undergraduate 3)

“This is the paper that I worked on.” (Post-qual Graduate 1)

“That’s for that conference paper.” (Pre-tenured Professor 2)

Here in the third example, the participant recognized an information item as a ‘conference paper’ because it is learned within the academic thought community. Thus, while this identification seems natural to other members of the academic thought community, non-members might not identify this information item as a ‘conference paper’. Similarly, in the case of the ‘time’ factor, often information items were identified based on the academic calendar, particularly by semester. For example participants said:

“It’s reading list for the summer vacation.” (Pre-qual Graduate 3)

“It’s the students in my class for the fall.” (Tenured Professor 3)

In the case of the ‘topic’ factor, participants often identified an information item based on the name of the research project or their research area, which showed an influence of the academia thought community. For example, participants said:

“It’s about Facebook use and civic and political engagement.” (Pre-qual Graduate 1)

“I do book culture and reading practices. It’s related to that.” (Post-qual Graduate 2)

In the case of the ‘related person’ factor, participants often identified an information item based on the relationships in academia. For instance, participants said:

“That’s my advisor’s edits for my Chapter Two of my dissertation.” (Post-qual Graduate 2)

“This next one is a student paper.” (Tenured Professor 4)

As shown above, the way participants recognize and identify an information item was often heavily influenced by their social background.

In the cases of the (4) Examination/Comparison stage and the (5) Selection/Modification/Creation stage, while assessing and adjusting mental gaps between old and new information items to make organizational decisions, ‘accessibility’, ‘affiliation’, ‘format’, ‘purpose’, ‘related person’, ‘time’, ‘topic’, and ‘use’ of an information item affected both stages. These factors are listed in the descending order of the frequency of occurrences, and are displayed in Table 19.

Table 19

Frequency of Occurrences of the Factors Impact on the (4) Examination/Comparison stage and the (5) Selection/Modification/Creation Stage

Rank	Factor	Occurrence
1	Purpose	133
2	Time	46
3	Accessibility	26
4	Topic	19
4	Related person	17
6	Use	16
7	Affiliation	14
8	Format	14

Among these factors, ‘accessibility’, ‘format’, and ‘use’ factors were not specifically associated with a particular thought community. For example, in the case of the ‘accessibility’ factor, when asked why they selected one of the categories, participants said:

“Because, it’s more relatively accessible.” (Early Undergraduate 2)

“It’s just easier to locate.” (Tenured Professor 2)

“And that way, I can just find it really quickly in my brain.” (Post-qual Graduate 2)

In addition, the ‘format’ of an information item did not show the influence of a particular thought community. For instance, while examining existing categories or selecting a certain category to organize an information item, participants said:

“I have a lot of categories in ‘Pictures’.” (Early Undergraduate 2)

“Because it’s a picture.” (Pre-qual Graduate 3)

Also, it was difficult to say that the ‘use’ factor was influenced by a particular thought community. While explaining why they decided to create, modify, or select a category, participants said:

“I create folders on my ‘Desktop’ for things that I’m working on.” (Tenured Professor 1)

“Because the old job is done so I possibly will not open the folder often than I used to do that.” (Pre-qual Graduate 1)

“‘Cause I don’t need it for a long time, so... I just figured that I’ll just save it there.” (Early Undergraduate 2)

As shown above, these factors were often not related to the particular thought community.

However, ‘affiliation’, ‘purpose’, ‘related person’, ‘time’, and ‘topic’ factors were often associated with the academia thought community. In the (4) Examination/Comparison stage and the (5) Selection/Modification/Creation stage, while assessing and adjusting mental gaps between old and new information items to make organizational decisions, participants often regarded the aspect(s) of an information item that are associated with the academic thought community as much more significant than other aspects, so that often the way participants grouped information items together or

separating certain information items from other information items were very similar to each other. In the case of the ‘affiliation’ factor, participants often grouped or separated information items based on the academic affiliation relevant to an information item, such as name of the university or department with which they were associated. For example, participants said:

“I have a folder for each of the universities I worked at.” (Tenured Professor 1)

“Then I started to just put everything related to ‘University 1’ here.” (Pre-qual Graduate 3)

“But the way I do it is I’ve put everything from ‘University 1’ into one folder.” (Post-qual Graduate 2)

“I’m in two departments, so two different departmental folders.” (Tenured Professor 4)

In the case of the ‘purpose’ factor, participants often made distinctions between categories primarily based on their academic tasks, such as classes they took or taught, and papers they wrote. For example, participants said:

“I arranged the whole structures based on conferences.” (Pre-tenured Professor 2)

“That folder is divided into various things like classes, writings, teaching, dissertation proposal.” (Post-qual Graduate 2)

“So I have a folder, of course, for each of my courses.” (Pre-qual Graduate 2)

“I have the ‘Classes’ folder.” (Early Undergraduate 3)

In a similar vein, the ‘related person’ factor was often influenced by the academia thought community. More specifically, participants sometimes grouped or separated information items based on the relationships in academia. For instance, participants said:

“There is folder for graduate students, graduate students I’ve had.” (Tenured Professor 1)

“If I have a graduate student who is particularly productive, I would have a folder named after him or her.” (Tenured Professor 2)

In addition, in the case of the ‘time’ factor, it was closely related to the academia thought community because often participants made distinctions between information items based on the academic calendar. The way participants made distinctions of ‘time’ were heavily influenced by academia, which eventually affected the way they made distinctions between information items. For example, participants said:

“I basically organize life by when I’m in school and when I’m not. So Christmas break from school was like, just a whole category of time when I was at home, and then as opposed to fall semester or spring semester when I’m here most of the time.... And then summer like, I’m at home again, so I basically organize it by when I’m in school or not in school. So I know in what point of life I was when I look back on it.” (Early Undergraduate 2)

“I have on my external hard drive as well as on my laptop, a ‘Spring Semester’ folder as well as ‘Fall Semester’, ‘Spring 2011’, et cetera.” (Late Undergraduate 1)

Last, in the case of the ‘topic’ factor, participants grouped or separated information items based on the ‘topic’ of an information item in a way that was often related to the academia thought community. For example, one participant said:

“Information technology is something that I have an interest in, so I have a folder in there called ‘IT’, so papers that I write in relation to IT get stored there.”

(Tenured Professor 2)

Similarly, when the researcher asked why a participant decided to save certain information item into a certain folder, he/she said:

“Same topic; same task. That’s the main reason.” (Post-qual Graduate 3)

As shown above, the aspects participants focused on while identifying an information item, which eventually affected the process of assessing and adjusting mental gaps between information items to either group them or separate them were greatly influenced by the academia thought community.

Common categories. To further analyze how the distinctions participants made were influenced by the academia thought community, the researcher analyzed categories in each folder structure that were recorded by the participants in the diary templates. As described in Section 4.4, in this research study, the researcher asked participants to record the path name of each category (i.e., folder structure) in the diary whenever they decided to save or organize an information item. An example of a folder structure follows:

“Documents/University 1/Spring 2012/Teaching/Slides”

In this research study, these folder structures were analyzed. The procedure used to analyze these categories is described below.

First, any redundant folder structures were deleted. Thus, if several information items were saved into a same category in the same folder structure, only one folder structure was counted and analyzed. Second, each folder in the folder structure was analyzed. For instance, in the case of this example folder structure,

“Documents/University 1/Spring 2012/Teaching/Slides”, all five folders, ‘Documents’, ‘University 1’, ‘Spring 2012’, ‘Teaching’, and ‘Slides’, were analyzed. In addition, in the case of any folder created based on two criteria, both criteria were counted and analyzed. For instance, in the case of the ‘Spring 2012’ folder, both the ‘semester’ criterion and ‘year’ criterion were counted separately. Third, the numbers of participants who used each of the categories were counted.

In total, from the folder structures which participants recorded in their diary templates over a week, 72 folder structures containing 174 folders were analyzed. Table 20 shows the names of the categories used by multiple participants in rank order. In addition, the number of participants who used each category is presented. In this table, all categories that were used by at least 5 participants were included.

Table 20
Categories Used By Multiple Participants

Rank	Category	Number of Participants
1	Year	10 (55.6%)
2	Course name (Class)	9 (50.0%)
3	Semester	7 (38.9%)
4	Documents	6 (33.3%)
4	Project name	6 (33.3%)
6	University name	5 (27.8%)
6	Conference name	5 (27.8%)
Total Participants		18

As shown in Table 20, the analysis of the data showed that participants had surprisingly similar categories. The category which was used by the highest number of participants was the ‘Year’ category, which was cited by 10 out of the 18 participants (55.6%). In addition, there were 25 different categories that used ‘Year’ as a criterion in grouping or separating information items. Further analysis of the data showed that among

those 25 categories, 8 of them (32.0%) were used together with the name of the conference (e.g., “Conference 1 2012”), 8 of them (32.0%) were used together with the semester name (e.g., “Spring 2012”), 5 of them (20%) were used together with certain tasks (e.g., “2010-Talk”), and 4 of them (16%) were used independently (e.g., “2012”). Particularly, among 5 categories that were used with certain tasks, 4 of them were academic tasks (e.g., “2011 Reviews”). In addition, 3 out of 4 of the ‘Year’ categories that were used independently were subordinate categories of a conference category (e.g., “/Conference 1/2012/”). Thus, although the ‘Year’ category is seemingly a neutral category which was not influenced by a particular thought community, analysis of the data showed that it actually was closely related to the academia thought community.

In the case of the ‘Class’ category, that it was used by the second highest number of participants showed that many participants grouped information items based on the class they were taking or teaching. Indeed, 9 out of 18 participants (50%) grouped information items based on the class, which showed a significant influence of the academia thought community on personal information organization. In this case, there were 20 different categories in which the name of the category was either ‘Class’ or specific course name (e.g., “Mediated Communication”).

The ‘Semester’ category was used by the third highest number of participants, which also revealed a significant influence of the academia thought community on the way participants organize their personal information items. As shown in Table 20, among 18 participants, 7 participants (38.9%) grouped or separated information items based on the semester (e.g., “Spring Semester”).

The 'Documents' category used by the fourth highest number of participants, but was not specifically relevant to the academia thought community. Actually, it was already created as a part of an Operation System (OS) on personal devices including desktop and laptops. Thus, some of the participants' folder structures started from this category, although it was not a category participants decided to create. Among 18 participants, 6 (30.0%) used this 'Documents' folder.

The 'Project name' category tied as the one used by the fourth highest number of participants, so that 6 out of 18 (30%) participants had a category by a research project name (e.g. "Social Q&A"), which reflected the influence of the academia thought community on the process of organizing personal information.

In a similar vein, the 'University name' category was used by 5 participants (27.8%), which indicates that almost one third of the participants grouped or separated information items based on the 'University' that was associated with information items (e.g., "University 1"). In addition, there were 15 different folder structures that included a category with a 'University name'.

The category of 'Conference name' (e.g., "Conference 2") was used by 5 participants (27.8%) among 18 participants.. There were 12 different folder structures that involved a category with a 'Conference name'.

As shown above, most participants had substantially similar categories, which showed the influence of the academia thought community on the process of organizing personal information. Participants often made distinctions between information items that would not be so obvious to non-members. Thus, the analysis of the categories showed

that participants often grouped or differentiated information items based on the particular thought styles of their thought communities, i.e., the academy.

5.2.6.2 The Process of Organizing Academic Versus Non-academic Information

The influence of the thought community was also found while analyzing the differences between the process of organizing information items that belong to the academic thought community and the non-academic thought community. As stated in Section 5.1.2, although all participants belonged to academia, they also belonged to thought communities other than academia. In the diary template, a number of participants recorded information items that belonged to the academia thought community as well as those that belonged to the non-academic thought community. Among 143 information items that were analyzed in this study, 112 of them belonged to the academia thought community, while 31 of them belonged to the non-academic thought community. Thus, examining the differences between the processes of organizing academic information items and non-academic information items seemed to be a good way of analyzing the influence of thought communities on the personal information organization process. Thus, information items that belonged to the academia thought community were grouped together and those that belonged to non-academic thought community were also separately grouped together. Then, the stages, actions, thoughts, decisions, and factors were comparatively analyzed.

Stages. To analyze the stages, the percentage of academic and non-academic information items that went through each stage of the process were counted and

comparatively analyzed. The stages of the process of organizing academic and non-academic personal information items are presented in Table 21.

Table 21
Stages of the Process of Organizing Academic and
Non-academic Information Items

Stage	Academic	Non-academic
1	110 (98.2%)	31 (100.0%)
2	99 (88.4%)	27 (87.0%)
3	32 (28.6%)	9 (29.0%)
4	61 (54.5%)	14 (45.2%)
5	96 (86.6%)	24 (77.4%)
6	96 (86.6%)	25 (80.6%)
Total Information Items	112	31

As shown in Table 21, all six stages were present in the process of organizing both academic and non-academic information items. As might be expected, for both types, almost all information items went through the (1) Initiation stage. In the case of the academic information items, among 112 information items, 110 of them went through the (1) Initiation stage, (98.2%) while 31 out of 31 non-academic information items (100.0%) went through this stage. In addition, most academic and non-academic information items went through the (2) Identification stage, (5) Selection/Modification/Creation stage, and the (6) Categorization stage, although the percentage of both academic and non-academic information items that went through each of these stages were slightly lower than those that went through the (1) Initiation stage. Fewer items went through the (4) Examination/Comparison stage, for both academic and non-academic information items. That is, 61 of 112 academic information items (54.5%) went through this stage, as did 14 of 31 non-academic information items (45.2%). The (3) Temporary Categorization stage had even fewer information items. In the case of academic information items, 32 of 112 of them went through the (3) Temporary

Categorization stage (28.6%), while 9 of 31 non-academic information items (29.0%) went through this stage. Thus, the stages participants went through, and the percentage of information items that went through each stage of the process were astoundingly similar, which means that the stages of the process were not particularly influenced by the thought community. This also indicated that the personal information organization model can be applied to the process of organizing personal information that belongs to thought communities other than academia.

Actions. To analyze the actions taken during the process of organizing academic and non-academic information, the percentages of occurrences for each action that was taken while organizing academic and non-academic information items was counted. In addition, the commonalities as well as differences were comparatively analyzed. The percentages of actions shown during the process of organizing academic and non-academic personal information items are presented in Table 22.

Table 22

Actions Taken During the Process of Organizing Academic and Non-academic Information Items

Stage	Code	Academic	Non-academic
1	Receive file	16 (14.3%)	5 (16.1%)
	Create file	21 (18.8%)	3 (9.7%)
	Save file	92 (82.1%)	28 (90.3%)
	Obtain file in another way	4 (3.6%)	1(3.2%)
3	Save file in a temporary location	32 (28.6%)	10 (32.3%)
5	Select a folder	90 (80.4%)	19 (61.3%)
	Modify previous folder	13(11.6%)	2 (6.5%)
	Create a new folder	9 (8.0%)	7 (22.6%)
6	Place file into a folder	92 (82.1%)	23 (74.2%)
Total Information Items		112	31

As shown in Table 22, the analysis of the data showed that there were several commonalities in the percentages of actions shown during the process of organizing both

academic and non-academic information items. However, there were some differences as well. In the (1) Initiation stage, a higher percentage of academic information items were created (18.8%) than non-academic information items (9.7%). However, the percentage of non-academic information items that was either received (16.1%) or saved (90.3%) was slightly higher than the percentage of academic information items that was either received (14.3%) or saved (82.1%).

In the (3) Temporary Categorization stage, a slightly higher percentage of non-academic information items (32.3%) were saved into a temporary location than academic information items (28.6%). In addition, in the (5) Selection/Modification/Creation stage, the percentage of information items that were organized by selecting one of the existing categories was higher for academic information items (80.4%) than that of non-academic information items (61.3%). Furthermore, the percentage of information items that were organized by modifying a previous category was higher for the academic information items (11.6%) than for non-academic information items (6.5%). However, for those information items that were organized by creating a new category, academic information showed a much lower percentage (8.0%) than non-academic information items (22.6%). During the process of organizing personal information, participants selected and modified existing folders more frequently when organizing academic information items than non-academic information items, while creating more new categories to organize non-academic information items than academic information items. This indicates that participants had more relevant categories which they already created for the academic information items than they did for non-academic information items. Thus, for the non-

academic information items, participants saved them into a temporary location or created a new category more than they did for the academic information items.

Thoughts. To comparatively analyze thoughts involved during the process of organizing academic and non-academic information items, the numbers of academic and non-academic information items that went through certain cognitive processes during the process of organizing information were counted. Then, similarities as well as differences were comparatively analyzed. The percentages of thoughts that were shown during the process of organizing academic and non-academic personal information items are presented in Table 23.

Table 23
Thoughts Shown During the Process of Organizing Academic and Non-academic Information Items

Stage	Code	Academic	Non-academic
1	Messy	0 (0.0%)	0 ⁴ (0.0%)
	Confusing	2 (1.8%)	0 ⁵ (0.0%)
2	Typify	74 (66.1%)	19 (61.3%)
4	Review existing folders	39 (34.8%)	12 (38.7%)
	Assess similarities and differences between new and existing files	45 (40.2%)	7 (22.6%)
5	Adjust mental gap between new and existing files	55 (49.1%)	12 (38.7%)
6	Clean	0 (0.0%)	0 ⁶ (0.0%)
Total Information Items		112	31

As shown in Table 23, there were some commonalities as well as differences in the percentage of information items that went through certain thoughts during the process of organizing information items, whether they belonged to the academia thought

⁴ This cognitive aspect was frequently mentioned by participants when they described about organization of information items which were not recorded in the diary because they were explaining about 1) their general organizing behaviors, 2) information items that were saved before they keep a diary for this study, or 3) what they will do in the future which were excluded when counting occurrences in the table.

⁵ See footnote 4.

⁶ See footnote 5.

community or a non-academia thought community. The percentages of information items that went through any specific cognitive process were similar for both types of information items.

However, in the (4) Examination/Comparison stage, while the percentage of reviewing existing categories was slightly higher for non-academic information items (38.7%) than for academic information items (34.8%), the percentage of assessing similarities and differences between new and existing information items was higher for academic information items (40.2%) than non-academic information items (22.6%). In other words, regardless of the number of categories participants had to examine and review to organize an information item, participants went through a higher percentage of the comparison thought process for the academic information items than for non-academic information items. As inferred in the comparative analysis of the behavioral aspect of the process, this indicates that participants had more relevant categories for academic information items. To be more specific, because there were multiple relevant categories in which information items could be organized, a higher percentage of information items went through the process of assessing similarities and differences between information items. It also means that the process of organizing academic information items might have been more difficult than for non-academic information items because participants had to assess similarities and differences between unorganized and organized information items for a greater percentage of academic information items than non-academic information items.

Decisions. The percentages of academic and non-academic information items that included certain decisions during the process of organizing information were comparatively analyzed, and the results are presented in Table 24.

Table 24

Decisions Made During the Process of Organizing Academic and Non-academic Information Items

Stage	Code	Academic	Non-academic
1	Save file	89 (79.5%)	24 (77.4%)
3	Delay decision making	19 (17.0%)	1 (3.2%)
	Re-categorize file	15 (13.4%)	3 (9.7%)
	Delete file	9 (8.0%)	4 (12.9%)
	Keep temporary categorization	7 (6.3%)	1 (3.2%)
5	Select a folder	83 (74.1%)	16 (51.6%)
	Modify previous folder	13 (11.6%)	2 (6.5%)
	Create a new folder	9 (8.0%)	5 (16.1%)
6	Delay decision making	33 (29.5%)	8 (25.8%)
	Re-categorize file	14 (12.5%)	1 (3.2%)
	Delete file	3 (2.7%)	5 (16.1%)
	Keep categorization	67 (59.8%)	16 (51.6%)
	Move file	1 (0.9%)	2 (6.5%)
Total Information Items		112	31

In the case of the decisions, like other aspects of the process, there were many similarities between academic and non-academic information items in the process of organization. However, there were some differences as well. On the one hand, in the (3) Temporary Categorization stage, the percentage of delayed decisions was much higher for the academic information items (17.0%) than for non-academic information items (3.2%). In addition, a higher percentage of academic information items were re-categorized (13.4%) than non-academic information (9.7%). The percentage of keeping decisions was also higher for the academic information items (6.3%) than non-academic information items (3.2%). In the (6) Categorization stage, the percentage of delayed decisions was higher for the academic information items (29.5%) than for non-academic

information items (25.8%). In particular, the percentage of information items that was re-categorized was much higher for academic information items (12.5%) than for non-academic information items (3.2%). In addition, a higher percentage of academic information items were kept in a category (59.8%) than non-academic information items (51.6%).

On the other hand, more non-academic information items were deleted both in the (3) Temporary Categorization stage (12.9%), and (6) Categorization stage (16.1%) than were academic information items in the (3) Temporary Categorization stage (8.0%) and (6) Categorization stage (2.7%).

To sum up, a higher percentage of academic information items involved a delayed decision, re-categorization decision, and keeping decision than non-academic information items, while a higher percentage of non-academic information items were decided to be deleted in both the (3) Temporary Categorization stage and the (6) Categorization stage. This indicates that the academic information items were used and kept for a longer period of time than non-academic information items. In addition, the decisions made during the process of organizing academic information were more complex than those of non-academic information items as shown when comparatively analyzing the cognitive aspect of the process of organizing academic and non-academic information items.

In addition, as shown in analyzing the behavioral aspect of the personal information organization process, in the (5) Selection/Modification/Creation stage, the percentage of information items which participants decided to organize by selecting one of the existing categories was higher for academic information items (74.1%) than that for non-academic information items (51.6%). Moreover, the percentage of information

items which participants decided to organize by modifying previous categories was much higher for the academic information items (11.6%) than for non-academic information items (6.5%). However, in the case of the percentage of information items which participants decided to organize by creating a new category, academic information items showed a lower percentage (8.0%) than non-academic information items (16.1%). Thus, as shown in the cross analysis of the behavioral as well as cognitive aspect of the process, it seems that participants had more relevant categories which they already created for the academic information items than they did for non-academic information items. Also, the decisions that were made during the process of organizing academic information items were more complex than for non-academic information items.

Factors. In the case of the factors that affected the process of organizing academic and non-academic information items, the top five factors were selected and comparatively analyzed; they are displayed in Table 25.

Table 25
Top Five Factors Influencing the Process of Organizing Academic and Non-academic Information Items

Rank	Academic		Non-academic	
	Factors	Occurrences	Factors	Occurrences
1	Purpose	105 (93.8%)	Purpose	24 (77.4%)
2	Use	87 (77.7%)	Use	22 (71.0%)
3	Time	42 (37.5%)	Accessibility	14 (45.2%)
4	Accessibility	41 (36.6%)	Format	10 (32.3%)
5	Topic	34 (30.4%)	Value	9 (29.0%)
Total		112	31	

As shown in Table 25, there were some common factors that heavily influenced the process. These factors were ‘purpose’, ‘use’, and ‘accessibility’. In particular, the

‘purpose’ and ‘use’ of an information item showed the highest number of occurrences for both academic and non-academic information items.

However, in the case of the academic information items, ‘time’ and ‘topic’ factors also heavily influenced the process, while these were not among the top five factors for the non-academic information items. Considering the fact that the ‘Semester’ category was used by 7 among 18 participants, and ‘Project name’ category was used by 6 among 18 participants, it was understandable that ‘time’ and ‘topic’ factors greatly affected the process of organizing academic information items.

In the case of organizing non-academic information items, ‘format’ and ‘value’ factors were two of the top five factors, although these were not included in the top five factors for the academic information items. For the ‘format’ factor, further analysis of the data showed that 18 of 31 non-academic information items (58.1%) were multimedia files (including picture, video, and music files), while only 3 out of 112 (2.7%) academic information items were multimedia files. For the academic information items, most of them (97.2%) were non-multimedia files (including text, PDF, Excel, and PowerPoint slides). It seems that the digital format of an information item worked as one of the influential factors during the process of organizing non-academic information items. In addition, the ‘value’ of an information item, which indicates a personal meaning attached to an information item, such as “interesting”, “funny”, “favorite”, and “cute”, influenced the process of organizing non-academic information items much more heavily than academic information items.

Depth of the organizational structure. To further analyze the differences, the depths of the organizational structure of the academic and non-academic information

items were analyzed. Again, in this process, redundant categories were deleted. The result from the analysis of the data showed that there indeed were differences in the depth of the folder structure for academic and non-academic information items. Participants had much more detailed and well-developed category structures for academic information items than for non-academic information items. Furthermore, the depth of the folder structure was compared by conducting an independent samples t-test, and the results showed that the depths of folder structure for academic information items ($M=2.73$, $SD=1.47$) were deeper than those of non-academic information items ($M=1.82$, $S=.80$), and there was a statistically significant difference between the depths of the organization structure for academic and non-academic information items: $t(66.43)=3.40$, $p < .01$. A simple box plot, which shows relative differences in distributions, is presented in Figure 13.

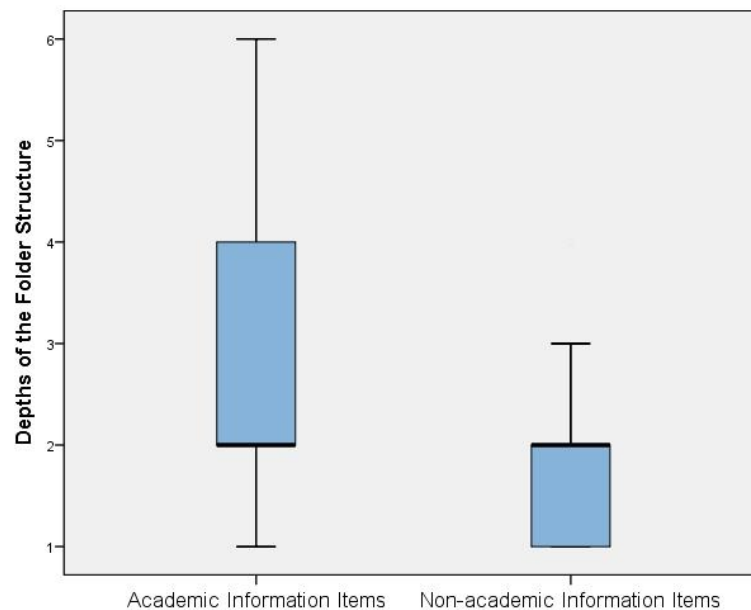


Figure 13. The Depths of the Folder Structure for Academic and Non-academic Information Items

As the boxplot shows, although the median values of the depths of categories for two different types of information items were almost the same, 50% of the depths of categories for academic information items were distributed between 2-4, while they were distributed between 1-2 for the categories for the non-academic information items. This shows that the participants had relatively more complex folder structures for the academic information items than for non-academic information items.

5.2.6.3 Influence of Professional Age

In this research study, one of the hypotheses was that the professional age of the participants, which is the length of the time participants spent as a member of a thought community, will influence the impact of the thought community on the process of organizing personal information. Thus, in this study, participants were recruited from three sub-groups in academia with different professional ages: (1) undergraduate students; (2) graduate students; and (3) professors. The average professional ages of participants in three different sub-groups are presented in Table 26.

Table 26
Average Professional Age of Participants in Different Sub-groups

Sub-group	Professional Age	Sub-sub group	Professional Age
Undergraduates	2.4	Early	1.5
		Late	3.3
Graduates	10.3	Pre-qualifying Exam	8.7
		Post-qualifying Exam	11.8
Professors	29.3	Pre-tenured	18.0
		Post-tenure	35.0

The influence of professional age was analyzed in various ways. First of all, the percentages of the academic and non-academic information items in each sub-group were analyzed. The result is displayed in Table 27.

Table 27
Number of Academic and Non-academic Information Items of Participants in Different Sub-groups

	Academic	Non-academic	Total
Undergraduate students	30 (61.2%)	19 (38.8%)	49
Graduate students	38 (84.4%)	7 (15.6%)	45
Professors	44 (89.8%)	5 (10.2%)	49

As shown in Table 27, all three sub-groups of participants had both academic and non-academic information items, and all three sub-groups had more academic information items than non-academic information items. However, there certainly were differences in the percentages of academic and non-academic information items. In the case of the undergraduate students, almost 60% of the information items were academic information items (61.2%), while about 40% were non-academic information items (38.8%). However, graduate students had a larger portion of academic information items (84.4%) rather than non-academic information items (15.6%) in their personal information items. In the case of the professors, an even higher percentage of personal information items were academic information items (89.8%), and non-academic information items accounted for only small portion of their personal information items (10.2%). Interestingly, while there was a large difference between undergraduate and graduate students regarding the percentages of academic and non-academic personal information items, there was a much smaller difference between graduate students and professors in the percentages of academic and non-academic personal information items. It seems that this is related to the commitment to the community. In Section 5.1.2, the analysis of the background questionnaire showed that undergraduate students were concerned with their professions after leaving academia, while graduate students and professors were more concerned with their professions as academics. In other words,

although all three sub-groups of participants were members of the academia thought community, undergraduate students were not as committed to the academia thought community as graduate students or professors were, because undergraduate students may or may not be members of the academia thought community in the future.

To sum up, the analysis showed that the longer participants spent time in academia, the higher percentages of academic information items they had. Thus, there were influences of professional age on the percentages of academic and non-academic information items participants had. In addition, in the percentages of academic and non-academic information items, the gap between the undergraduate students and graduate students was bigger than that of graduate students and professors.

Common categories. The influence of professional age was also examined by comparatively analyzing the folder structures of three sub-groups. As described in Section 5.2.6.1, the researcher investigated the folders in each folder structure that were recorded by the participants in the diary templates. In addition, to analyze the influence of professional age, categories that were used by multiple participants in each sub-group were examined and analyzed. In Table 28, categories that were used by multiple participants in each sub-group are presented together with the number of participants.

Table 28
Categories Used by Multiple Participants in Different Sub-groups

Category Name	Undergraduates	Graduates	Professors
Year	3 (50.0%)	3 (50.0%)	4 (66.7%)
Course name (Class)	3 (50.0%)	3 (50.0%)	3 (50.0%)
Semester	4 (66.7%)		2 (33.3%)
Project name		4 (66.7%)	2 (33.3%)
Conference name		2 (33.3%)	3 (50.0%)
Documents		3 (50.0%)	2 (33.3%)
University name		2 (33.3%)	2 (33.3%)
Poster		2 (33.3%)	
Picture			2 (33.3%)
Teaching			2 (33.3%)
Personal			2 (33.3%)
Department name			2 (33.3%)
Participant name			2 (33.3%)
Review			2 (33.3%)
Total Participants	6	6	6

Interestingly, as shown in Table 28, each sub-group of participants had several categories that were used by multiple participants, which indicates that participants lumped or split their personal information items in similar ways. In particular, the ‘Year’ and ‘Course name’ categories were used by multiple participants in all three sub-groups. Moreover, these two categories were used by more than 50% of participants in each group, which shows a profound influence of the academia thought community on the participants’ personal information organization process.

However, there were clear differences among sub-groups in the number of common categories that were used by multiple participants. In the case of the undergraduate students, there were three common categories that were used by multiple participants. That is to say, other than these three categories, each participant had unique categories. However, in the case of the graduate students, there were seven categories

which were used by multiple participants. This shows that the graduate students, who spent more time in academia, grouped or separated their personal information items in more similar ways than undergraduate students. This was even more evident among professors. In the case of the professors, there were 13 common categories that were used by multiple participants, which indicates that the way professors grouped or separated information items were surprisingly similar. Thus, as the number of year participants spend as an academic increased, the way they organize their personal information became much more similar, which reveals the influence of professional age on the process of organizing personal information, and shows that categories are often socially learned.

The influence of professional age was also examined by analyzing the commonalities and differences of the process of organizing information items by participants in three sub-groups. In particular, the stages, actions, thoughts, decisions, and factors that were involved during the process of organizing personal information by participants in different professional ages were comparatively analyzed.

Stages. The percentages of information items that went through each stage of the personal information organization process while the three sub-groups of participants organized their personal information items were comparatively analyzed, and are presented in Table 29.

Table 29
Stages of the Information Organization Process of Participants in Different Sub-groups

Stage	Undergraduates	Graduates	Professors
1	49 (100.0%)	44 (97.8%)	47 (95.9%)
2	39 (79.6%)	43 (95.6%)	46 (93.9%)
3	15 (30.6%)	8 (17.8%)	18 (36.7%)
4	28 (57.1%)	23 (51.1%)	23 (46.9%)
5	41 (83.7%)	38 (84.4%)	41 (83.7%)
6	42 (85.7%)	38 (84.4%)	41 (83.7%)
Total Information Items	49	45	49

As shown in Table 29, when organizing personal information, participants in all three sub-groups went through all six stages of the process. Particularly, in the case of the (1) Initiation stage, the (4) Examination/Comparison stage, the (5) Selection/Modification/Creation stage, and the (6) Categorization stage, there was no large difference in the percentages of information items that went through each of these stages among participants in the three groups. Almost all information items went through the (1) Initiation stage, so that the percentages ranged from 95.9% to 100%. In the case of the (4) Examination/Comparison stage, about half of the information items that were organized by three sub-groups of participants went through this stage, with the percentages ranging from 46.9% to 57.1%. In addition, in the (5) Selection/Modification/Creation and the (6) Categorization stage, most of the personal information items that were organized by all three sub-groups of participants went through this stage, so that the percentages ranged from 83.7% to 85.7%. It can be seen that the percentages of information items that went through each of the stages were surprisingly similar among participants in different professional ages.

However, the percentage of information items that went through the (2) Identification stage during the process of organizing information by undergraduate

students (79.6%) was somewhat lower than that of graduate students (95.6%) or professors (93.9%). In addition, the percentage of information items that went through the (3) Temporary Categorization stage during the process of organizing information by graduate students (17.8%) was somewhat lower than that of undergraduate students (30.6%) or professors (36.7%). However, in these cases, the percentages were neither increasing nor decreasing according to the professional ages of participants so that it is difficult to say that this shows the influence of professional age on the stages of personal information organization process. Thus, it seems that there was no major difference in the stages that participants in different professional ages went through during the process of organizing personal information.

Actions. The actions taken during the process of organizing each personal information item by the three sub-groups of participants in different professional ages were comparatively analyzed, and are presented in Table 30.

Table 30
Actions Taken During the Information Organization Process of Participants in Different Sub-groups

Stage	Code	Undergraduates	Graduates	Professors
1	Receive file	1 (2.0%)	11 (24.4%)	9 (18.4%)
	Create file	11 (22.4%)	8 (17.8%)	5 (10.2%)
	Save file	40 (81.6%)	37 (82.2%)	43 (87.8%)
	Obtain file in another way	0 (0.0%)	4 (8.9%)	1 (2.0%)
3	Save file in a temporary location	16 (32.7%)	8 (17.8%)	18 (36.7%)
5	Select a folder	35 (71.4%)	34 (75.6%)	40 (81.6%)
	Modify previous folder	13 (26.5%)	1 (2.2%)	1 (2.0%)
	Create a new folder	5 (10.2%)	6 (13.3%)	5 (10.2%)
6	Place file into a folder	37 (75.5%)	35 (77.8%)	43 (87.8%)
Total Information Items		49	45	49

As shown in Table 30, there were similarities as well as differences. In the (1) Initiation stage, the percentages of information items that had been saved were similar for

all three sub-groups, ranging from 81.6% to 87.8%. However, in the case of the undergraduate students, participants received a much lower percentage of information items (2.0%) when compared with that of graduate students (24.4%) or professors (18.4%). In addition, undergraduate student participants created a higher percentage of information items (22.4%) than did graduate student participants (17.8%) or professor participants (10.2%). This could mean that undergraduate students collaborate with others less than graduate students or professors. However, it could also be related to when the data were collected (i.e., at the beginning of the semester or session) so that they collaborated with others less than they usually do.

In the (3) Temporary Categorization stage, graduate students saved a lower percentage of information items in a temporary location (17.8%) than did undergraduate students (32.7%) or professors (36.7%). In the (5) Selection/Modification/Creation stage, the percentages of information items that had been organized by selecting one of the existing categories were similar among all three sub-groups of participants, ranging from 71.4% to 81.6%. The percentages of information items that had been organized by creating a new category were also similar among all three sub-groups, ranging from 10.2% to 13.3%. However, in the case of the percentage of information items that had been organized by modifying existing categories, the percentage of information items that were organized by undergraduate students was much higher (26.5%) than that of graduate students (2.2%) or professors (2.0%). The further analysis of the data showed that this is because undergraduate participants often modified existing folder structures by creating subordinate categories based on class, week, or exam to differentiate certain information items from other information items. Thus, while the analysis of the data showed some

traits of certain sub-groups, there were no significant differences in the actions shown during the process of organizing personal information among the three sub-groups of participants.

Thoughts. The percentages of information items that went through certain cognitive processes during the process of organizing personal information by the three sub-groups were investigated while comparatively analyzing their similarities and differences. The percentages of thoughts that were shown during the process of organizing each of the information items by three sub-groups of participants are displayed in Table 31.

Table 31
Thoughts Shown During the Information Organization Process of Participants in Different Sub-groups

Stage	Code	Undergraduates	Graduates	Professors
1	Messy	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Confusing	2 (4.1%)	0 (0.0%)	0 (0.0%)
2	Typify	34 (69.4%)	28 (62.2%)	31 (63.3%)
4	Review existing folders	14 (28.6%)	15 (33.3%)	22 (44.9%)
	Assess similarities and differences between new and existing files	19 (38.8%)	18 (40.0%)	15 (30.6%)
5	Adjust mental gap between new and existing files	19 (38.8%)	28 (62.2%)	20 (40.8%)
6	Clean	0 (0.0%)	0 (0.0%)	0 (0.0%)
Total Information Items		49	45	49

As shown in Table 31, there were some commonalities as well as differences among the three sub-groups of participants regarding the percentages of information items that went through certain thought processes during the process of organizing information items. There were more similarities than differences, in that the percentages of information items that went through most of the thought processes were similar among all three sub-groups of participants.

However, in the (4) Examination/Comparison stage, the percentages of folders that were reviewed by participants while they decided where to organize information items increased along with the professional ages of sub-group of participants. The undergraduate student participants reviewed existing categories when they were organizing 28.6% of information items. The graduate students examined existing folders while organizing 33.3% of their personal information items, which was slightly higher than that of undergraduate students. In the case of the professors, existing categories were reviewed while organizing 44.9% of their information items. In short, the longer participants had spent time in academia, the more they reviewed existing categories, which indicates that those who spent more time in the thought community had more relevant categories to examine when deciding where to organize an information item. Thus, there was an influence of professional age on the cognitive aspects of the personal information organization process.

In the case of the (5) Selection/Modification/Creation stage, the percentage of information items that went through the process of adjusting mental gaps between new and existing information items was somewhat higher for the graduate participants sub-group (62.2%) than that of the undergraduate participants sub-group (38.8%) or professor participants sub-group (40.8%). However, the percentages were neither increasing nor decreasing according to the professional ages of participants, so that it is difficult to say if this was an influence of professional age.

Decisions. The percentages of information items that involved certain decisions while three sub-groups of participants organized their personal information items were comparatively analyzed. The results from the analysis are presented in Table 32.

Table 32
Decisions Made During the Information Organization Process of Participants in Different Sub-groups

Stage	Code	Undergraduates	Graduates	Professors
1	Save file	36 (73.5%)	39 (86.7%)	38 (77.6%)
3	Delay decision making	5 (10.2%)	1 (2.2%)	14 (28.6%)
	Re-categorize file	7 (14.3%)	1 (2.2%)	10 (20.4%)
	Delete file	4 (8.2%)	7 (15.6%)	2 (4.1%)
	Keep temporary categorization	3 (6.1%)	0 (0.0%)	5 (10.2%)
5	Select a folder	34 (69.4%)	27 (60.0%)	38 (77.6%)
	Modify previous folder	13 (26.5%)	1 (2.2%)	1 (2.0%)
	Create a new folder	4 (8.2%)	6 (13.3%)	4 (8.2%)
6	Delay decision making	10 (20.4%)	5 (11.1%)	26 (53.1%)
	Re-categorize file	10 (20.4%)	1 (2.2%)	4 (8.2%)
	Delete file	7 (14.3%)	0 (0.0%)	1 (2.0%)
	Keep categorization	23 (46.9%)	35 (77.8%)	25 (51.0%)
	Move file	1 (2.0%)	0 (0.0%)	2 (4.1%)
Total Information Items		49	45	49

In the case of the decisions, there were more differences than similarities.

However, in most cases, it was hard to say that those differences are the result of the influence of professional ages of participants. In the (3) Temporary Categorization stage, on the one hand, the percentages of information items that involved delayed decisions (2.2%), and the percentage of information items that were decided to be re-categorized (2.2%) were much lower for the graduate students than undergraduate students (whose decisions involved 10.2% of delayed decisions and 14.3% of re-categorization decisions), or professors (whose decisions involved 28.6% of delayed decisions and 20.4% of re-categorization decisions). On the other hand, for the graduate participants, the percentage of information items that were decided to be deleted (15.6%) in this stage was much higher than that of undergraduate student participants (8.2%) or professor participants (4.1%). Thus, although there were some differences in the percentages of information items that involved certain decisions during the process of organizing personal

information, the percentages did not show either ascending or descending order when arranged based on the professional ages of the participants, so that it was difficult to say that there was an influence of professional age on the decisions made during the process.

In a similar vein, in the (6) Categorization stage, the percentages of some of the decisions made by graduate students such as delayed decisions (11.1%), re-categorization decisions (2.2%), and deleting decisions (0.0%) were much lower than the percentages of those decisions made by undergraduate participants (20.4%, 20.4%, and 14.3%, respectively) or professors (53.1%, 8.2%, 2.0%, respectively). However, the percentage of keeping a decision made by graduate students was much higher (77.8%) than that of undergraduate participants (46.9%) or professors (51.0%).

In the case of the (5) Selection/Modification/Creation stage, as shown when analyzing the behavioral aspects of the process, the percentage of information items that were decided to be organized by modifying existing categories (26.5%) was much higher for undergraduate student participants than for graduate student participants (2.2%) or professor participants (2.0%). Thus, although there were some differences in the percentages of certain decisions that were made during the process of organizing personal information among the three sub-groups of participants, it did not specifically show the influence of professional ages of participants. Thus, there was no significant difference among participants in different professional ages in the decisions made during the process of personal information organization.

Factors. The factors that heavily influenced the process of organizing personal information items by the three sub-groups were examined. Among 19 factors that affected the process, the top five factors that influenced the process of organizing

personal information by each sub-group were selected and comparatively analyzed. The top five factors are presented in Table 33.

Table 33
Top Five Factors Impact on the Information Organization Process of Participants in Different Sub-groups

Undergraduates		Graduates		Professors	
Factors	Occurrences	Factors	Occurrences	Factors	Occurrences
Purpose	42 (85.7%)	Purpose	41 (91.1%)	Purpose	46 (93.9%)
Use	33 (67.3%)	Use	34 (75.6%)	Use	42 (85.7%)
Accessibility	22 (44.9%)	Topic	20 (44.4%)	Accessibility	19 (38.8%)
Time	20 (40.8%)	Type	17 (37.8%)	Related person	18 (36.7%)
Necessity of differentiation	11 (22.4%)	Accessibility	14 (31.1%)	Topic	16 (32.7%)
Total Information Items	49		45		49

As shown in Table 33, on the one hand, there were some common factors:

‘purpose’, ‘use’, and ‘accessibility’. Especially, the ‘purpose’ and ‘use’ of an information items showed the highest number of occurrences for all three sub-groups of participants.

On the other hand, different factors affected the process of personal information organization of participants in different professional ages. In the case of the undergraduate students, the ‘time’ and ‘necessity of differentiation’ factors were included as two of the top five factors, while these were not included in the top five factors for the graduate students or professors. Regarding the fact that three categories that were used by more than 50% of the undergraduate student participants were ‘Year’, ‘Course name’, and ‘Semester’ as shown in Table 28, it is understandable that the ‘time’ factor was selected as one of the top five factors. In addition, participants often decided to organize an information item by modifying an existing category by creating subordinate categories to differentiate certain information items from other information items (as shown in Table

30 as well as Table 32), so the fact that the ‘necessity of differentiation’ factor was selected as one of the top five factors that affected the undergraduate participants’ process of organizing personal information was not surprising.

In the case of the graduate students, the ‘type’ factor was included as one of the top five factors, while it was not one of the top five factors that influenced the process of organizing personal information of undergraduate student participants or professor participants. In the case of the professor participants, the ‘related person’ factor was the unique factor selected as one of the top five. Further analysis of the data showed that this was because professors collaborated more often with other people than did other sub-groups of participants.

Interestingly, in the case of graduate students and professors, among the top five factors that affected the personal information organization process, four of them (80%) were the same. Thus, as the professional age increased, factors that influenced the process became more similar, which shows the role of professional age on the impact of a thought community on the personal information organization process.

Case study. The influence of professional age on the process of organizing personal information was not only analyzed by investigating the number of information items that involved certain stages, actions, thoughts, decisions, and factors as well as folder structures, but also by examining individual participants’ organizing behaviors. In this section, one of the early undergraduate students and one of the tenured professors’ categories for academic information items are illustrated to better explain the influence of professional age on the process of organizing personal information.

The ‘Early Undergraduate 1’ who is one of the undergraduate participants in this research study spent 1.5 years in academia. In the first interview, when the researcher asked which role the participant regarded as the most important source of identity for him/her among three roles he/she mentioned, the participant answered:

“Definitely school” (Early Undergraduate 1)

However, in organizing personal information, he/she had one big category in which he/she saved all of the personal information items related to academic tasks. During the first interview, the participant said:

“For classes, um, I put all of them in ‘My Document’.” (Early Undergraduate 1)

This participant said that he/she created this category and had been using it since his/her freshman year. Thus, when the researcher asked the participant whether it was easy for him/her to make a decision about where to save an information item, the participant answered:

“Um, yeah, because I just decided to put everything under one file (folder) so that’s what that was.” (Early Undergraduate 1)

However, during the first interview, the participant mentioned that he/she recently realized that this organization made it difficult to find needed information items. For instance, the participant said:

“I realized lot of my stuff was very confusing and everything was like all over the place in the ‘Documents’.” (Early Undergraduate 1)

One of the reasons for the difficulty was the number of information items. In other words, it was because the number of information kept increasing. For example, the participant said:

“Maybe I should categorize my notes and everything and not just put it all in ‘Documents’ ’cause maybe that will be too confusing ’cause I have so much information on the ‘Documents’.” (Early Undergraduate 1)

More importantly, it was because the participant began to take multiple courses in his/her department which included similar information items. As a consequence, the number of courses the participant took in his/her department increased so that several files in this category had similar names. This started to make it difficult to differentiate certain information items from others, which eventually hindered finding needed information items. During the interview, the participant described this difficulty as well as the need for revising his/her categories. The participant said:

“Maybe I’ll create different folder, like for each class, so that I wouldn’t get so confused because sometimes, I took a few Communication classes, so when it says, ‘Comm’ or ‘Communication something’, then like, I get confused because, I think it’s for this class, but really it’s for different class, so, maybe I would re-organize it.” (Early Undergraduate 1)

In consequence, during the few weeks between the first and the second interview, the participant actually re-organized his/her organizational structure. More specifically, within ‘My Documents’ category, the participant created subordinate categories based on the classes he/she took or was taking. Thus, in the second interview, the participant showed re-organized information items, and said:

“I tried to group everything that was for one class into one folder in ‘My Documents’.” (Early Undergraduate 1)

This case is a good example which shows how professional age affects the process of organizing personal information. As more time was spent in a particular thought community, not only did the number of information items grow, but the undergraduate participant also started to develop a more refined organization structure for his/her information items relevant to that thought community. In this process, the participant started to make distinctions between information items which non-members would not make. Although the participant eventually found his/her initial category inappropriate in organizing his/her personal information items, it actually worked for the participant during the period when he/she spent less time in academia. However, it is not difficult to imagine that this undergraduate participant's initial organization would not be appropriate for those who had spent a longer time as members of academia.

The next case is another good supporting example for this assertion. The 'Tenured Professor 1' is one of the professor participants in this research study who had spent 46 years as a member of the academia thought community. Just as the undergraduate participant said, this participant also answered that he/she regarded his/her role as an academic as a vital source of identity. The participant said:

“Being an academic is the most important source of identity for me.” (Tenured Professor 1)

However, unlike the previously described undergraduate participant who used one big category to organize his/her information items relevant to academia, this professor participant had a complex yet very refined and well-developed organizational structure for his/her information items, based on various tasks he/she performed in academia. The participant had a folder for each university he/she worked at, and in each university

category, he/she had a number of sub categories, sub-sub categories, and sub-sub-sub categories which showed that he/she made fine distinctions between information items relevant to the academia. The participant said:

“Over time, I created a very complex set of hierarchical file systems of things that arrive at my desk that I know about, that I have to deal with, it’s part of my profession. And so I have names for all sorts of things. And these are things that I consistently get. So I have very standard set of folders that I go into and store things on regular basis. So it’s really quite organized. When you start out as a new faculty member, probably you don’t have that. But then you start keep developing over time.” (Tenured Professor 1)

It is important to note that the participant also mentioned that his/her organizational structure was not something that was done at once; instead, it was developed gradually over time as the years he/she spent in the academic thought community increased.

As illustrated by the personal information organization of the two participants who spent relatively shorter and longer periods of time in academia, the longer people spent time as a member of a thought community, the more they identified, lumped, and split personal information items by specific tasks as a member of the thought community. Although both groups said that their job as an academic was the most important source of their identities in their daily lives, the undergraduate students initially did not particularly made distinctions between information items as an academic. By contrast, participants who spent a longer time as academics had sophisticated organizational structures which contained fine distinctions they made as an academic. Hence, the analysis of the data shows that as the amount of time a person spends in a certain thought community

increases, people make more specific distinctions between things as a member of a particular thought community.

5.2.7 The PIOP Model

RQ7. Does the PIOP (Personal Information Organization Process) model explain the process of organizing personal information?

RQ8. Do participants always go through certain stages, actions, thoughts, decisions, and factors during the process?

As described in Section 4.6.2.2, the researcher developed a PIOP model for each information item. However, among 235 diary entries (i.e., information items), some of them were basically redundant in terms of investigating the process of organizing personal information. To be more specific, some of them were different versions of the same file, or multiple articles that were downloaded from the web for the same course or the same paper. In these cases, instead of giving full descriptions about the process, participants often said:

“This is all basically the same.” (Early Undergraduate 1)

“It’s just a different version of that file” (Pre-tenured Professor 2)

In these cases, information items were not developed into a process model because 1) it was redundant and 2) most of the time, the verbal descriptions of the process in these cases were incomplete. Thus, among 235 information items, 143 information items were used in developing a model. In each model, the researcher wrote the information ID and the name of the information file on the left top of the model, and wrote the type of the

specific thought community of the information item in the left bottom of the box that shows the stages of the process. Then, the researcher identified stages which the participant went through during the process of organizing that information item. The sequence of the process was expressed by using rectangles (stages) and arrows (sequence). In addition, any actions, thoughts, decisions, and factors involved in organizing that specific information item were presented under each stage of the process. An example model of one of the information items is presented in Figure 14.

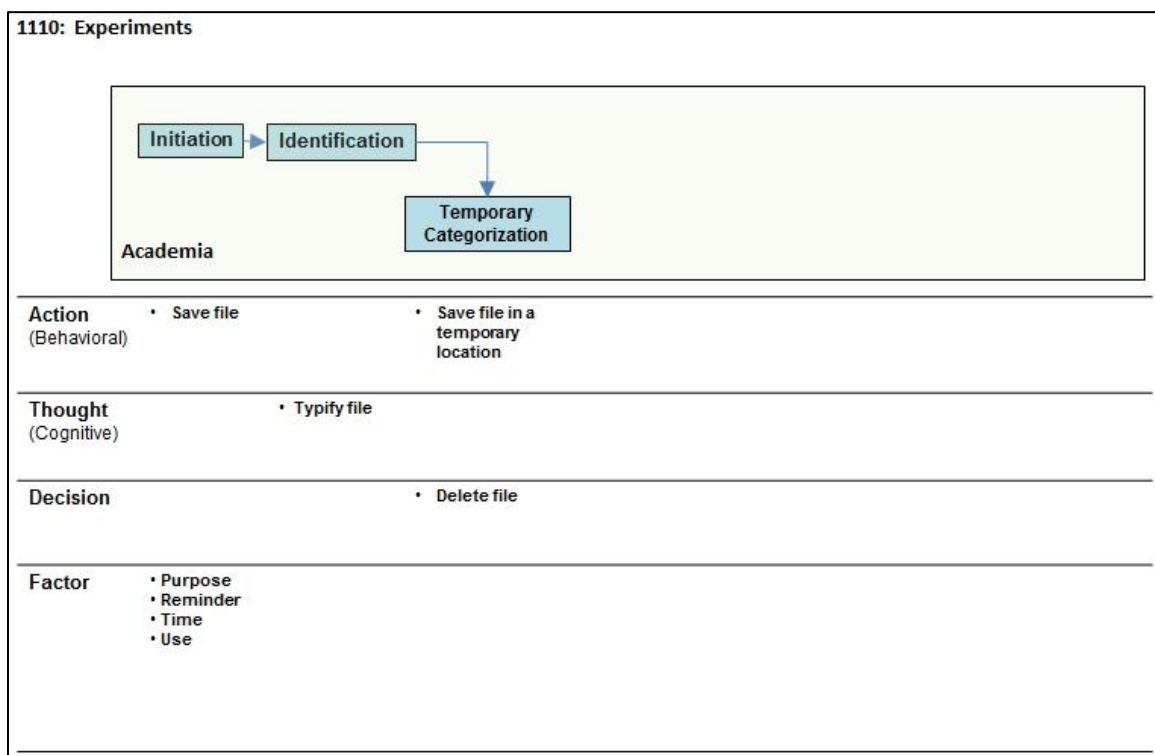


Figure 14. An Example of the PIOP Model for an Information Item

In the cases when participants described future decisions or possible processes, the rectangles as well as arrows were presented in a dotted line. In addition, the actions, thoughts, decisions, and factors mentioned by participants as possible future behaviors

that they might make are noted in gray letters. This can be found in one of the example models of the information item displayed in Figure 15.

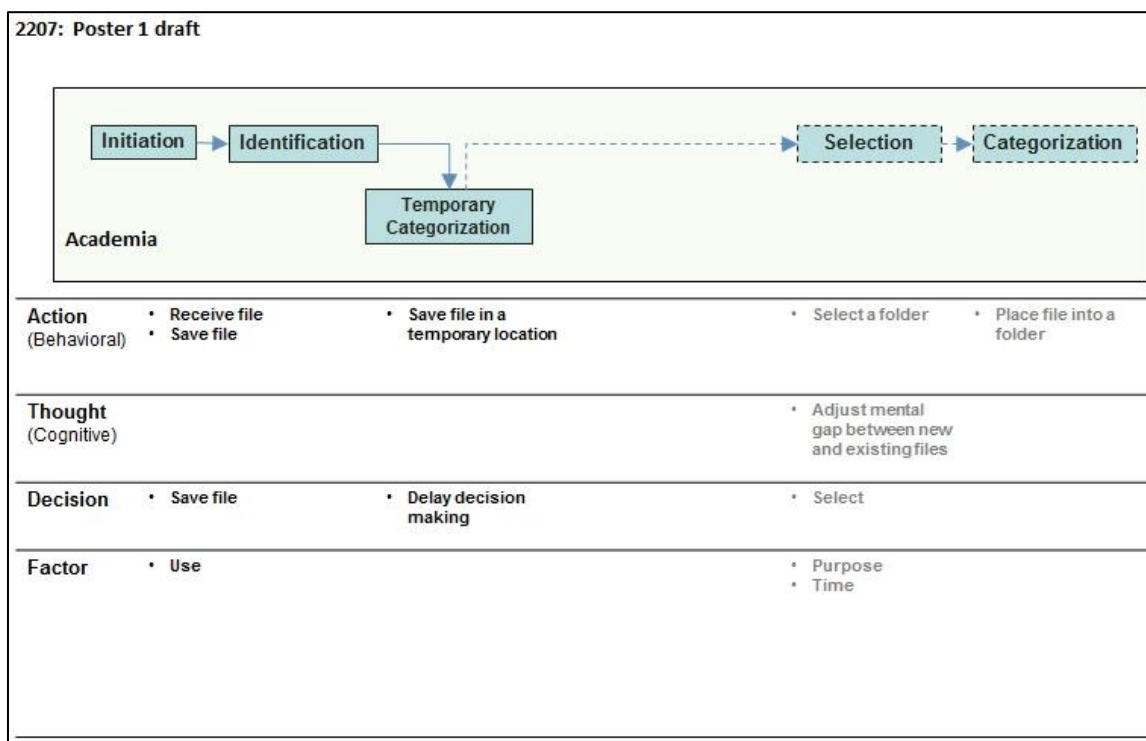


Figure 15. An Example of the PIOP Model for an Information Item with Future Decisions

After developing a model for each of the 143 information items, the researcher developed an integrated version of the model which best shows the process of organizing personal information. The final version of the PIOP model developed in this research study is presented in Figure 16.

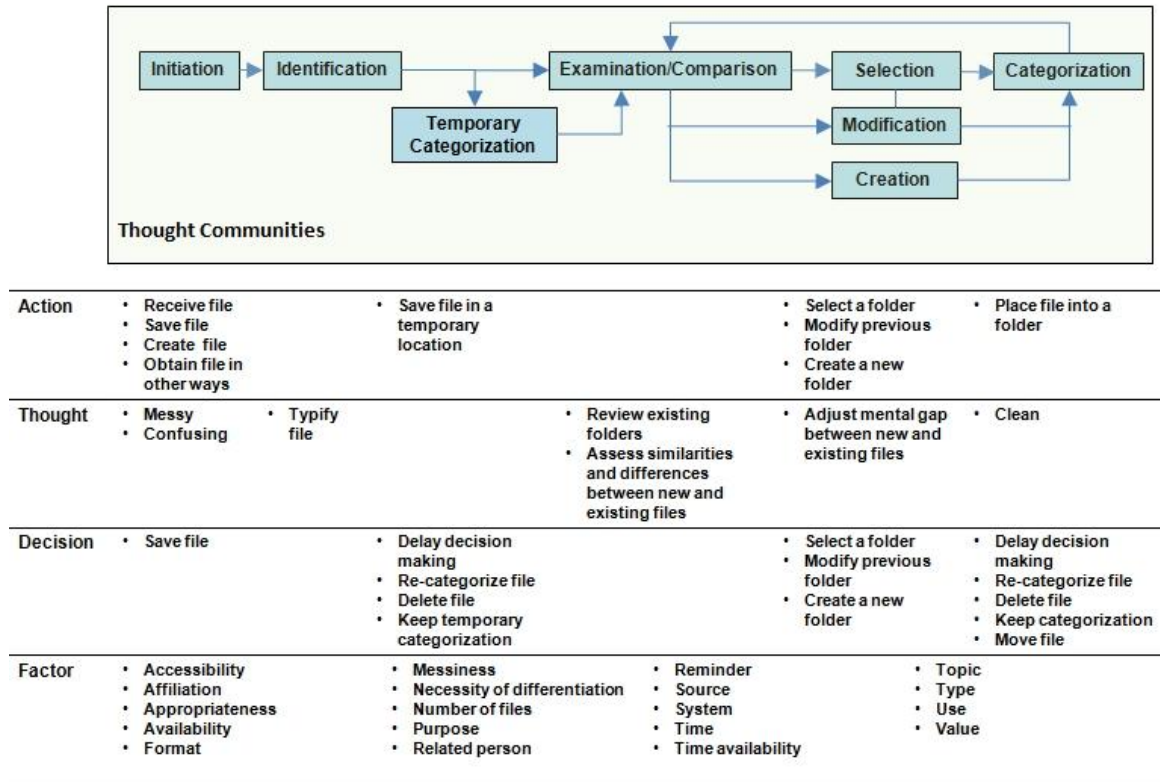


Figure 16. The Personal Information Organizing Process (PIOP) Model

This is a newly developed model in this study. As shown above, this model explains the process of organizing personal information by holistically showing different stages participants went through, the sequence of the stages, different actions participants took in each stage of the process, thoughts that were involved in each stage of the process, decisions participants made in each stage of the process, and various factors that influenced the process of organizing personal information. More specifically, the process of organizing personal information in digital form had six stages which involved different actions, thoughts, decisions, and a variety of factors that affected the process. In addition, the whole process was heavily influenced by individuals' thought communities.

However, it is important to note that not everyone went through all six stages, all actions, thoughts, and decisions, or was affected by all factors presented in this model. As previously mentioned, the model aims to show ranges of behaviors, rather than typifications of behaviors. Thus, the specific elements in the model are examples of what can happen during the process, rather than comprehensive elements.

CHAPTER 6 DISCUSSION

In this research study, the researcher investigated the personal information organization process in a systematic way and developed a model that helps explain what happens during the process of organizing personal information in a holistic way. In particular, the stages, actions, thoughts, decisions, and factors involved during the process of organizing personal information are investigated. In addition, the social aspect of the process, which is the influence of a thought community on the process, is examined. Some of the interesting findings of this research study, and their relationship to the previous studies, are highlighted and discussed in this section.

6.1 The PIOP Model

6.1.1 Stages: Temporary Categorization and Types of Information Items

In this study, the researcher initially developed a new model that shows the personal information organization process based on the researcher's analysis of the literature. Then this initial model was confirmed, modified, and extended while analyzing the empirical data. The final model includes six different stages, which are (1) Initiation, (2) Identification, (3) Temporary Categorization, (4) Examination/Comparison, (5) Selection/Modification/Creation, and (6) Categorization.

Among the six stages of the process of organizing personal information, the (3) Temporary Categorization stage was the stage which was not included in the initial model, but newly found while analyzing actual data. More specifically, the researcher found that often participants delayed decisions and saved information items in a

temporary location for a short period of time without examining or selecting existing categories in their folder structures.

It seems that this stage is closely related to the organization of ‘action information’, which is identified by Cole (1982) as an information item that is in current use or will be used in the near future, as well as ‘personal work file’, which is identified as an information item relevant to people’s ongoing work (p. 60). Cole also stated that this ‘action information’ is usually placed on desks or the floor so that participants can immediately access it; and similarly, ‘personal work file’ is placed on the stack of files, in cupboards, filing cabinets, or in the participants’ desk, which are all part of the immediate office environment (p. 60). Although Cole investigated the organization of paper-based information, this ‘action information’ and ‘personal work file’, which were placed where participants can easily access them, seemed closely related to the (3) Temporary Categorization stage of the process in this research study, in which participants saved information items into a temporary location such as ‘Desktop’ or ‘Downloads’ when there was an ongoing use of the file, or when the files need to be accessed frequently or quickly. In addition, this stage also seemed relevant to ‘working information’, which is identified by Barreau and Nardi (1995) as an information item that is directly related to the participants’ current work and frequently used by the participants (p.41). Especially in the case of Barreau and Nardi’s study, the authors examined the organization of digital forms of personal information rather than paper-based forms, and reported that this ‘working information’ was organized either in its own folder or location on the digital desktop.

As shown above, the approaches used in previous studies and this research study were different. Previous studies intended to identify different types of personal information items, while this research study attempted to identify different stages of the process. However, the findings of the previous studies and this study support each other by showing that regardless of the format of the personal information item, participants tend to save information items that are in use or need to be accessed frequently in an easily accessible place. In addition, the findings also revealed that this location is often different from the categories which are part of the well developed organizational structures. These results also imply that people usually have not one but multiple organizational systems for different ‘uses’ of an information item, and not all personal information goes through the same stages of the organization process. In this study, all types of information items and the processes they went through were integrated to the PIOP model. However, it would be interesting to identify different types of personal information items based on the ‘use’ of information item, and explore the process of organizing each type of information item.

6.1.2 Actions: Reluctance to Change the Structure of Organization

The findings of this research study showed that participants took different actions in different stages of the process of organizing personal information. After obtaining information items in the (1) Initiation stage, participants sometimes saved files in a temporary location in the (3) Temporary Categorization stage. Other times, in the (5) Selection/Modification/Creation stage, participants selected a folder, modified a previous folder, or created a new folder to organize information items that they obtained or saved

in a temporary location. Lastly, in the (6) Categorization stage, participants placed files into a folder.

Interestingly, the analysis of the data indicated that in the (5) Selection/Modification/Creation stage, participants tended to keep their organization structure by selecting one of the existing categories rather than modifying the existing structure. The results show that almost 80% of the actions taken in the (5) Selection/Modification/Creation stage were selecting one of the existing folders, while only about 10% to 11% of the actions taken in this stage were modifying existing folders or creating a new folder.

This finding supports Douglas' (1978) assertion that people are reluctant to revise their existing categories; usually, people accept not only those items that easily fit into the existing categories, but also somewhat ambiguous items by treating them as if they are one of the members of the category. However, if the items are so different from other items that they cannot fit into any categories, people often ignore or distort them so that they do not have to modify their structure of assumptions which they have established and kept with confidence over time (p. 36). Indeed, in this research study, the analysis of the empirical data shows that during the process of organizing personal information, participants obviously had their own organizational structure which they developed, and tried to keep the classification structure rather than changing it. In addition, some of the participants were more rigid or flexible than others in maintaining the developed organizational structure.

It would be interesting to investigate how personal information organizational structure develops over time, and to further analyze the characteristics of people who are more rigid or flexible in keeping their organizational structure than others are.

6.1.3 Thoughts: Messiness, Typifications, and Category's Blurred Boundary

This study investigated the cognitive aspect of the process of organizing personal information, about which little is known. In particular, it found that participants felt cognitively 'messy' when information was not organized. Initially, the researcher assumed that the cognitive status of the (1) Initiation stage would be 'undefined'. However, the analysis of the data shows that when information items were not organized, participants felt it was messy, which often made them initiate the organization process. Thus, it helped to understand one of the motivations of organization (i.e., cognitive messiness).

This cognitive 'messiness' corresponded with Douglas' (1978) explanation about 'dirt', which is a 'matter out of place' that needs to be put back in the place where it belongs (Zerubavel, 1991). Moreover, in this study, participants mentioned that they felt it was cognitively clean when information was placed into a category, which indicates that cognitively, personal information organization involves a process of cleaning (i.e., eliminating the 'dirt'). Thus, it was interesting to find how Douglas' work applies to the process of organizing personal information.

In addition, the findings of this study showed that before making organizational decisions, typifications occurred. At the early stage of the process of organizing personal information, participants identified an information item by focusing on a certain aspect of

it while disregarding other aspects of it so that they could typify it as ‘something’. Then, they made organizational decisions (i.e., where to categorize an information item). This finding confirms previous studies which stated that typification is a prerequisite for classification (Berger & Luckmann, 1966; Cooper, 2004; Zerubavel, 1991). This study also shows that this can be applied to the process of organizing personal information.

In addition, the analysis of the data reveals that participants certainly reviewed categories which were relevant to the unorganized information item, and assessed similarities as well as differences between the new information item and existing information items in the relevant categories to make decisions about where to categorize an information item. This result confirms the context theory of classification (Medin & Schaffer, 1978), which asserted that people organize information items by comparing new information items with their previous information items, and the classification judgment derives from stored information.

In addition, the findings show that participants adjusted the mental gap between new and existing information items to decide where to organize an information item, which indicates that information items can be organized into several overlapping categories. This finding corresponds with previous studies which stated that there is no clear boundary to a category or concrete defining properties that determine the membership of the category (Lounsbury, 1956; Rosch, 1978; Wittgenstein, 1953). However, although this nature of the category was examined by various scholars in different fields, those studies mostly examined this characteristic of category either conceptually or by investigating the categorization of objects in the natural world. Therefore, by analyzing the empirical data, the results from this study further confirm it

by showing that this characteristic of categories applies to the classification of personal information items.

6.1.4 Decisions: Graded Structure of Categories

This study investigated the decisions that were made during the process of organizing personal information, and found that participants made a variety of decisions in different stages of the process. Regardless of the type of decisions they made, participants mentioned that some of the information items were easier to make decisions about, while others were more difficult. This result indicates that the structure of a category is graded, meaning that some members of a category are better examples than others. When an information item was a better member of a category, participants found it easier to organize than other information items.

Although the focus of her work was not on the organization of information, in the prototype theory, Rosch (1978) defined ‘prototypes’ of categories as “the clearest cases of category membership defined operationally by people’s judgments of goodness of membership in the category” (p. 36). Zadeh’s (1965) fuzzy set theory also argued that there is a continuum of grades of membership in a category. In addition, this graded structure of categories have been studied and supported by many other researchers (Berlin & Kay, 1969; Ekman, Friesen, & Ellsworth, 1972; Rosch, Simpson, & Miller, 1976) as introduced in Section 2.1.

This study showed that this characteristic of a category can be applied to the organization of personal information in digital forms. Also, previous studies help to understand why categorization of certain information items was easier or more difficult

than other information items, and why categorization cannot be done simply and perfectly.

6.1.5 Factors: Comparative Analysis

The findings of this research study show that a variety of factors affect the process of organizing personal information in digital form. In this study, 19 factors that influenced the personal information organization process were identified. Among previous studies relevant to this research study, Kwasnik's (1989) study, which investigated various factors that influenced professors' organizational decision of classification of paper-based personal information items in their offices, is the study that is most closely related to this research study. As described in Section 2.3, Kwasnik identified 34 factors that influenced the organizational decisions, which she called 'dimensions' of classificatory decisions in her study (p. 82). Among those factors identified in Kwasnik's study, and among the 19 factors that were identified in this research study, the top ten factors are displayed in Table 34.

Table 34
Top Ten Factors Impact on Personal Information Organization

Kwasnik's Study	This Research Study
Form	Purpose
Use/Purpose	Use
Topic	Accessibility
Time	Time
Locate	Topic
Access	Related person
Arrangement	Format
Source	Type
Related to me	Number of files
Physical attributes	Source

Although only the top ten factors are presented, Table 34 shows that there were some differences in the way factors were labeled in each of the studies. For example, while the ‘purpose’ and ‘use’ factors were used as separate factors in this research study, they were combined in Kwasnik’s study. In addition, sometimes different terminology was used to represent certain factors, such as ‘format’ in this study and ‘form’ in Kwasnik’s study. Also, there were some unique factors that were identified only in one of the studies, such as ‘types’ factor in the current study and ‘locate’ factor in Kwasnik’s study. In addition, it is important to note that while Kwasnik’s study identified various dimensions mentioned by participants when they were describing how they made organizational decisions for their personal information items, this study explored the factors that affected the process of personal information organization, including not only organizational decisions made during the process, but also actions and thoughts involved during the process.

However, regardless of these differences, Table 34 shows that the top factors that were identified in both studies are surprisingly similar. Among top ten factors, seven factors were almost the same, which indicates that the most influential factors identified in both studies are very similar. In both studies ‘purpose’ and ‘use’ factors were identified as one of the most influential factors. It was the second highest ranked in Kwasnik’s study, while it was the first and second highest ranked in terms of their occurrences in this research study. As a matter of fact, there have been other studies which have reported that the ‘purpose’ of an object influences the construction of categories or perceived similarity of objects (Barsalou, 1983; Ratneshwar, Barsalou, Pechmann, & Moore, 2001). The results from this research study confirm that the

‘purpose’ and ‘use’ of personal information items greatly influence personal information organization.

In addition, Table 34 shows that in both Kwasnik’s study and this research study, ‘accessibility’, ‘topic’, ‘format’, ‘source’, and ‘time’ that is associated with an information item heavily influenced the organization of personal information. In fact, although the exact numbers of occurrences or rankings of each factor were not provided, Barreau’s (1995) study, which examined whether Kwasnik’s findings could be applied to the personal information management systems in an electronic environment, reported that ‘form’, ‘locate’, ‘topic’, ‘circumstance’, and ‘use/purpose’ factors were mentioned most often when participants organized their personal information in a digital form (p.333). In the case of Barreau’s (2008) study, which was conducted to examine whether advances in technology influenced participants’ personal information management, she reported that participants organized their information based on ‘task’, ‘topic’, ‘provenance’, and ‘form’ (p. 315-316). Thus, although there were some differences, factors that heavily influence personal information organization seem quite consistent in all of the studies. These influential factors included ‘purpose/use’, ‘topic’, and ‘format’ of an information item. In addition, ‘time’ and ‘source’ also worked as important factors that affected the organization.

In Kwasnik’s study, she categorized the 34 factors into seven broad groups, which were (1) situation attributes, (2) document attributes, (3) disposition, (4) order/scheme, (5) time, (6) value, and (7) cognitive state. Among the top ten factors presented in Table 34, ‘access’, ‘related to me’, ‘source’, and ‘use/purpose’ factors belong to the (1) situation attributes; ‘form’, ‘topic’, and ‘physical attributes’ factors belong to the (2)

document attributes; 'locate' factor belongs to (3) disposition; 'arrangement' factor belongs to the (4) order/scheme; and 'time' factor belongs to (5) time. Thus, Kwasnik found that not only document factors but also various situational factors affected the organizational decisions (p. 393).

When the top ten factors that were identified in this research study were grouped into the seven broad categories of Kwasnik's study, 'purpose', 'use', 'accessibility', 'related person', and 'source' factors could be grouped into (1) situation attributes; 'topic', 'format', and 'type' factors were grouped into (2) document attributes; 'number of files' factor were categorized into (4) order/scheme; and 'time' was categorized into (5) time. This study thus supports Kwasnik's argument that not only document factors but also situational factors influence the organization of personal information, by showing that her findings can be applied to the process of organizing personal information. Barreau's studies, which were conducted in 1995 and 2008, also reported that both document attributes and situation attributes primarily affected the personal information organization decisions. How the top ten factors identified in both studies can be grouped into seven categories is presented in Table 35.

Table 35

Top Ten Factors Impact on Personal Information Organization Grouped into Categories

Category	Kwasnik's Study	This Research Study
(1) Situation Attributes	Access	Accessibility
	Related to me	Purpose
	Source	Use
	Use/Purpose	Related person
		Source
(2) Document Attributes	Form	Format
	Topic	Topic
	Physical attributes	Type
(3) Disposition	Locate	
(4) Order/Scheme	Arrangement	Number of files
(5) Time	Time	Time

The findings of this research study showed that a variety of factors affected the personal information organization. The factors that were identified as influential in previous studies on organization of personal information, including 'purpose/use', 'topic', 'format', 'time', and 'source' mostly appeared as top factors in this research study. Therefore, this study confirms and extends previous studies by showing that influential factors are consistent, and those factors that include the situational factors heavily influence not only personal information organization decisions but also actions and thoughts that are involved during the process.

6.1.6 Model: Dynamic and Holistic Process of the Information Organization

At the beginning of the study, it was assumed that the organization process is ongoing, and cannot be completed at certain point, but rather keeps changing over time. To investigate those changes, the second interview was included in the study design. In fact, the findings of this research study show that not only when participants organize an

information item, but also after participants placed an information item into a certain category, they make various decisions so that organizational structure keeps changing over time. The analysis of the data shows that although sometimes the temporary categorization was maintained for a certain period of time, often those information items that were saved in a temporary location got deleted or re-categorized into the folder structure, especially when there was no ongoing use of the file, or when participants wanted to keep it for possible future use. In addition, although participants often selected one of the categories in the existing folder structure which resulted in strengthening the existing folder structure, sometimes participants created new categories, particularly when there was a new task, when there were a number of related files, or when participants anticipated having more files. In addition, participants sometimes modified the existing categories by splitting them into subcategories or merging them into a superordinate category. In the case of splitting, participants created subordinate categories when they felt certain information items needed to be differentiated from other information items. In the case of merging, participants often merged folders into a superordinate category when there were no ongoing uses of the files, particularly after a certain period of time (such as at the end of the project, semester, or year). Moreover, even after information items were placed into a category in the folder structure, participants not only kept them but also re-categorized them, moved them, or deleted them. In the case of re-categorization, it was usually made when participants realized that the categorization was inappropriate, either because they made wrong decisions or the value or use of information items had changed. Information items were often moved to another personal device when participants wanted to make backups or keep them for a

long period of time. When there was no future use for the files or folders, participants deleted them. Thus, the organizational structure was not static but dynamic. No organizational structure could be thought as permanent.

In fact, in the PIOP model that was developed in this study, the arrows which demonstrate the sequence of the process do not exhibit a simple linear direction that starts from the (1) Initiation stage and ends in the (6) Categorization stage. They show multiple directions which sometimes go back to the previous stage. This reflects the dynamic process of personal information organization. Although such the dynamic characteristics have been studied by other researchers (Ravasio, Schär, & Krueger, 2004; Whittaker, 2011), it was not the primary focus of the study. However, in this study, the dynamics of the personal information organization process have been thoroughly investigated and reflected in the PIOP model, which made it easier to understand the process. Thus this study makes a contribution to the field. It would be interesting to revisit participants in this study after a certain period of time and examine further changes that have been made to their personal information organizational structure.

The PIOP model developed in this study shows the ranges of actions, thoughts, decisions, and factors involved during the process of organizing personal information, so that it helps when understanding the process holistically. As addressed in Section 2.2, there have been some interesting studies which examined the process of certain information behaviors, especially those which examined information seeking behavior (Bates, 1989; Belkin, 1980; Kuhlthau, 1991; Savolainen, 1995; Wilson, 1997). However, studies that specifically explored the process of organizing information have been missing. In addition, most previous studies in personal information organization focused

on certain aspects of the process rather than the process as a whole (Bälter, 1997; Barreau 1995, 2008; Barreau & Nardi, 1995; Bergman, Whittaker, Sanderson, Nachmias, & Ramamoorthy, 2010; Boardman & Sasse, 2004; Case, 1991; Cole, 1982; Fisher, Bruxh, Gleve, & Smith, 2006; Gonçalves & Jorge, 2003; Henderson & Srinivasan, 2009; Kwasnik, 1989, 1991; Malone, 1983; Whittaker & Sidner, 1996). Thus, while some of the findings of this study were already investigated by other researchers, previous studies could not show how each aspect of the process interacts with each other during the process of organizing personal information.

Kuhlthau's (1983) study explored the process of information searching while investigating affective, cognitive, and physical aspects of the process, and it is one of the studies most relevant to this current one. In particular, she developed a model, the ISP (Information Search Process) model, which not only exhibits common patterns people engage in during the process of information seeking, but also shows how "people experience the information search process holistically, with an interplay of thoughts, feelings, and actions" (Kuhlthau, 2005 , p. 230). However, Kuhlthau examined the process of information searching rather than organizing, so that the holistic model that explains the process of information organizing has been missing.

Therefore, this study not only extends previous findings but also makes an unique contribution to the field by developing a model that provides an integrated view of the process organization of personal information, which shows how each aspect of the process interacts with the others during the process of organizing personal information.

6.2 Social Aspect of the Process of Organizing Personal Information

6.2.1 The Influence of the Thought Community: Mind Structure

In this study, cognitive sociology is used as a theoretical framework, so that the process of organizing personal information is interpreted in this framework. To be more specific, as stated in Section 3.1, this research study viewed the personal information organization process not only as a universal process and an individual process, but also as a social process which is influenced by people's social environment. The cognitive sociological perspective provided further explanation about the process, which would have not been possible without using this framework. This section addresses and highlights how society, and especially thought communities of participants influenced the process.

In Section 5.2.6, the social aspect of the process has been analyzed. The factors that affected the (2) Identification stage, the (4) Examination/Comparison stage, and the (5) Selection/Modification/Creation stage were further analyzed to investigate the impact of the academia thought community. In addition, the process of organizing academic and non-academic information items was comparatively analyzed. Moreover, the influence of professional ages on the process of organizing personal information was analyzed. The findings of this study show that the participants' information organization process was indeed influenced by their main thought community, i.e., academia. The ways participants identified an information item, and the ways in which they assessed and adjusted the mental gaps between information items, were heavily influenced by the academia thought community. During the process, among various aspects of an information item, participants tended to focus on the aspect that is associated with

academia (such as academic tasks, academic calendar, academic affiliation, or relationships in academia). In essence, often academic aspects of the information items outweighed other aspects of the information, so that participants often recognized and identified an information item based on those aspects, and made distinctions between information items which would not be clear to non-members. In fact, the analysis of folders showed that participants had remarkably similar categories, and those categories used by multiple participants were primarily associated with academia, which indicates the influence of their thought community.

The findings of this study support the view of cognitive sociology which asserts that the reasons why members of particular thought communities lump or split things in similar ways, and often make distinctions between things which non-members would fail to notice, is because things are often considered as socially similar or different as people classify things as members of particular thought communities (Zerubavel, 1996).

The comparative analysis of the process of organizing academic and non-academic personal information showed that there were no significant differences in the stages or in the ranges of actions, thoughts, decisions, and factors that were shown in the PIOP model. This indicates that the model can be applied to the process of organizing personal information that belongs to thought communities other than academia. In particular, all six stages were present in the process of organizing both academic and non-academic personal information.

However, in the case of the behavioral aspects of the process, participants organized a higher percentage of academic information items either by selecting or modifying one of the existing categories than they did for non-academic information

items. They also organized a higher percentage of non-academic information items either by saving in a temporary location or creating a new category than they did for academic information. This was also confirmed in the percentage of decisions participants made. The percentage of information items which participants decided to organize by selecting one of the existing categories was higher for academic information items than for non-academic information items. It was also higher than the percentage of information items which participants decided to organize by modifying a previous category. Also, just as in the behavioral aspect, the percentage of information items which participants decided to organize by creating a new category was much higher for non-academic information items than for academic information items.

In addition, in the case of cognitive aspects of the process, participants assessed similarities and differences between new and existing information items more frequently when they were organizing academic information items than non-academic information items. Moreover, in the case of decisions, a higher percentage of academic information items that were saved in a temporary location were re-categorized into one of the existing folders, while a higher percentage of non-academic information items were deleted. Likewise, a much higher percentage of academic information items that were placed into a category was kept or re-categorized than were non-academic information items, while a higher percentage of non-academic information items were deleted.

The fact that a higher percentage of academic information items involved the process of comparing new and existing information items, and selecting or modifying one of the existing categories than did non-academic information items, while a higher percentage of non-academic information items involved saving them in a temporary

location or creating a new folder than did academic information items indicates that participants had more relevant and similar categories for the academic information items than for non-academic information items. This also means that participants had more developed categories for academic information items which entailed more and finer levels of distinctions. The fact that participants had more developed and deeper organizational structures for academic information items was also confirmed when the depths of the folder structure for academic and non-academic information items were analyzed.

Therefore, the comparative analysis shows that the process of organizing academic information is more complex than that of non-academic information in this thought community. In addition, the fact that a higher percentage of academic information items were either kept or re-categorized than non-academic information items, while a higher percentage of non-academic information items were deleted showed that more academic information items were used and kept for a longer period of time than non-academic information items.

Here are some possible explanations for the reason why participants had more well-developed and finer categories for academic information items. One of them is that academia was the participants' primary thought community, so that participants had more information items relevant to their primary thought community, which eventually resulted in more elaborate categorization. If this explanation is true, people in a particular thought community will have more developed categorization for information items that belong to their primary thought community, regardless of the type of thought community. Another possible explanation is that this is one of the characteristics inherent of either

people in academia or of academic information items. In fact, Barreau's (2008) study, which investigated personal information organization of managers in a company or a government agency rather than academics, reported that participants left most information items in a temporary location (p. 311). In this research study, a higher percentage of non-academic information items was saved in a temporary location without folders than academic information items. Thus, it may be possible either that academic information items require finer categorization, or that academics tend to make finer distinctions. To test out whether one of the explanations or both of the explanations provide an accurate view of the process, further investigation will be needed. However, both possible explanations reveal that the thought community of participants affects the process of organizing personal information.

It will be interesting to investigate the process of organizing personal information of people in a thought community other than academia to further examine the social aspect of the process.

It is important to note that this does not mean that individual differences do not exist. In fact, the analyses of the data showed that there certainly are individual differences. For example, some participants preferred to have a well-developed folder structure, saying that when files are not organized, it was difficult to find things. One of the participants said:

"I would rather have hierarchies than huge long list. It's easier to search."

(Tenured Professor 1)

Other participants preferred to have less developed folder structures, saying that it made them easier to access information items and saved time. For example, one of the participants said:

“I like to see it visible. I like the visuality of seeing everything at the same level.”

(Tenured Professor 4)

Zerubavel (1997) also stated that “People in any given social environment are clearly not all cognitive clones of one another, which suggests that the way we think is by no means determined totally by society. Each of us is a member of more than just one thought community and therefore inhabits several different social worlds. As a result, we each have a rather wide “cognitive repertoire” and often think somewhat differently in different social contexts” (p.17)

However, the thought community of these participants indeed affected the process of organizing personal information, so that participants often perceived similarities and differences among information items, and made distinctions between them in similar ways. Both Brekhus (2010) and Zerubavel (1997) stated that because thought communities shape people’s perspectives and the ways people perceive the world around them, eventually it makes members of the communities perceive things similarly. In other words, society constructs the way people think, including the way people perceive similarities and differences between things, which influenced the participants in constructing their own socio-mental lenses for viewing reality. Then, because the way participants’ organized their personal information reflected this mental structure (i.e., how they viewed and organized reality) it resulted in similar organizational structures

among participants. In fact, during the interview, when the researcher asked participants why they organized information items in certain ways, one of the participants said:

“I really divide my mind that way.” (Post-qual Graduate 2)

Thus, by using the cognitive sociological perspective, this research study highlighted the influence of thought communities on the process of organizing personal information, which often can be overlooked.

6.2.2 The Influence of Professional Age: Socialization

Whether the professional age of participants affected the impact of the thought community on the process of organization or not was investigated to further examine the influence of thought community on the process of organizing personal information. Thus, three sub-groups of participants in different professional ages were comparatively analyzed.

In the case of the comparative analysis of the process of organizing personal information by the three sub-groups of participants in different professional ages, there were no significant differences in the stages or in the ranges of actions, thoughts, decisions, and factors that were shown in the PIOP model. This indicates that the model can be applied to the process of organizing personal information of participants by different professional ages.

However, the analysis of the folder structures revealed that the professional age of participants heavily increased the influence of the thought community so that the longer participants spent time as a member of a thought community, the more they organized their personal information in similar ways. This was most evident in the number of

common categories that were created by sub-groups of participants. The analysis of the data shows that the number of identical categories used by multiple participants in each group increased in accordance with the professional age. The analysis of the data shows that the number of common categories graduate students had was more than twice than that of undergraduate students, while the number of common categories professors had was more than three times than that of undergraduate students. In the case of professors, 13 common categories were used by multiple professors, which shows that professors had substantially similar organization for their personal information items. Moreover, those identical categories were primarily associated with the academia thought community. In addition, the in-depth analysis of the personal information organization of 2 participants who spent relatively shorter and longer periods of time in academia showed the influence of professional age; the longer people spent time as a member of a thought community, the more they identified and made finer distinctions among personal information items as a member of a thought community.

This finding that participants perceived information items more similarly as academics when their professional ages increased can be interpreted as socialization into a particular thought community. In fact, the individual is not born as a member of society, but rather he/she becomes the member of society in a specific social structure in a temporal sequence (Berger & Luckmann, 1966). As the amount time people spent in a particular thought community increases, people start to learn the unique thought style of that thought community (either consciously or unconsciously) so that just like socialization, as time goes by people become more accustomed to that thought community's mental lenses through which they see the world. Although the context was

different, Cooper (2004), who investigated cognitive categories for library information in a group of children in kindergarten through grade 4, also found and reported that children move from a more personal understanding toward a more sociocultural understanding of information in the library. Thus, just as children learn how to view and classify information based on sociocultural understanding about information, people learn to view and organize information items based on the socio-mental lenses of a particular thought community. Zerubavel (1997) stated that:

“We likewise learn to see things as similar to or different from one another. After all, whenever we classify things, we always regard only some of the differences among them as significant and ignore others as negligible and therefore irrelevant, yet which differences are considered significant is something we learn, and ignoring those that “make no difference” involves tacit social pressure to disregard them despite the fact that we do notice them, just as we learn that in order to find a book in a bookstore we must attend to the first letters of its author’s last name while ignoring the color of its cover. Separating the relevant from the irrelevant, as we shall see, is not just a logical but also a normative matter.” (p.13-14)

Thus, the findings of this study further support the social aspect of personal information organization by revealing that when organizing personal information, professional ages of participants significantly influenced the impact of thought community on the process so that the more participants spent time as a member of thought community, the more they organized information items in similar ways (i.e., as academics who reflected the thought style of academia). This implies that when

conducting comparative analysis of the process of personal information organization of people in different thought communities, the thought style of the members will be more evident with people who spent a longer time as a member of thought community than who spent less time.

CHAPTER 7 CONCLUSION

7.1 Answering the Research Questions

RQ1. What are the stages of the personal information organization process?

The process of personal information organization involved six stages, which were (1) Initiation, (2) Identification, (3) Temporary Categorization, (4) Examination/Comparison, (5) Selection/Modification/Creation, and (6) Categorization.

RQ2. What actions do people take during the process of organizing personal information?

During the process of organizing personal information, participants showed ranges of behaviors. In the (1) Initiation stage of the process, participants received, saved, or created files. Sometimes, participants obtained files in other ways. In the (3) Temporary categorization stage, participants saved files in a temporary location. In the (5) Selection/Modification/Creation stage, participants either selected one of the existing folders, modified one of the existing folders, or created a new folder. In the (6) Categorization stage, participants placed files into a folder.

RQ3. What cognitive process do people go through during the process of organizing personal information?

The process of organizing personal information involved different cognitive processes. In the (1) Initiation stage, participants felt it was messy or confusing when files were not organized. In the (2) identification stage, participants typified

files. In the (3) Examination/Comparison stage, participants reviewed existing folders or/and assessed similarities and differences between new and existing files in relevant folders. In the (5) Selection/Modification/Creation stage, participants finally adjusted the mental gap between new and existing files. In the (6) Categorization stage, when files were organized, participants felt it was cognitively clean.

RQ4. What decisions do people make during the process of organizing personal information?

Participants made various decisions during the process of organizing personal information. In the (1) Initiation stage, participants decided whether to save files or not. In the (3) Temporary Categorization stage, participants sometimes delayed decision making, re-categorized files into the folder structure, deleted files, or kept temporary categorization. In the (5) Selection/Modification/Creation stage, participants sometimes decided to select one of the existing folders, modify one of the existing folders, or create a new folder. In the (6) Categorization stage, participants sometimes delayed decision making, deleted files, kept categorization, moved files into another personal device, or re-categorized files either by placing them into another category or re-structuring the folder structure.

RQ5. What factors influence the process of organizing information?

There were various factors that affected the process of organizing personal information. In this study, 19 factors that influenced the process of organizing personal information were identified. These factors included (1) accessibility, (2) affiliation, (3) appropriateness, (4) availability, (5) format, (6) messiness, (7)

necessity of differentiation, (8) number of files, (9) purpose, (10) related person, (11) reminder, (12) source, (13) system, (14) time, (15) time availability, (16) topic, (17) type, (18) use, and (19) value.

RQ6. How do individuals' thought communities influence the process of organizing personal information?

The thought communities of participants indeed influenced the process of organizing personal information. The way participants identify files, group or differentiate files were not only remarkably similar but also was primarily associated with academia. In addition, the professional ages of participants greatly affected the impact of the thought community on the process, so that the influence of the thought community on the process was more evident in the organization of those who spent a longer time in the thought community.

RQ7. Do participants always go through certain stages, actions, thoughts, decisions, and factors during the process?

The PIOP model aimed to show ranges of behaviors that were involved during the process of organizing personal information rather than typifications of behaviors. Thus, not all participants went through all the stages, actions, thoughts, or decisions that are presented in the model.

7.2 Implications

7.2.1 Implications for Research

There have been a number of research studies which found numerous critical and insightful findings in understanding information organization. However, a research study

that specifically explored the process of organizing personal information had been missing. In particular, social influence on the personal information organization process had been not investigated.

In this research study, the researcher systematically investigated the process of organizing personal information. Particularly, based on the analysis of the literature as well as the empirical data, an integrated model (i.e., the PIOP model) which explains the process of personal information organization has been developed. This model provides a holistic view of the process by identifying stages of the process as well as actions, thoughts, decisions, and various factors that were involved in each stage. The findings of this study show how each aspect of the process is connected to and interacts with the others during the process of organizing personal information.

The findings of this research advance our knowledge and understanding about people's information organization process, of which little is currently known. Because this research is an interdisciplinary endeavor that is related to multiple academic fields including information science, library science, philosophy, psychology, sociology, cognitive science, human-computer interaction, and computer science, it lays out an empirical foundation for further study of any information organizing behaviors. This research also deepens our understanding of how people manage their information.

Besides the PIOP model, the influence of thought communities on the process has been examined by using cognitive sociology as the theoretical framework which extends our understanding about the process. Taking a cognitive sociological perspective, which views an individual's classification behavior as the product of socialization, (Brekhus, 2007), makes a unique contribution to the field by providing a new explanation of the

social aspects of personal information organization. In the past, classification of personal information was considered as an individual behavior. However, cognitive sociology provides another way of thinking about how and why people organize their personal information in certain ways. This research study shows the influence of thought communities on the process, and demonstrates that the process of personal information organization is not just an individual process. However, although this study is related to the idea of domain analysis, it differs from domain analysis. While domain analyses do not consider individuality, this study allowed for both social and individual aspects of the process.

7.2.2 Implications for Practice

Effective personal information management, especially personal information organization, is directly related to an individual's everyday life in pursuing different goals, and performing diverse roles and responsibilities. Currently, there are various devices, interfaces, and tools that are designed to support an individual's managing or organizing their personal information in digital forms. However, to design tools that effectively support people, system designers need to understand and reflect individuals' behaviors and preferences. Otherwise, although they were designed to facilitate an individual's personal information organization, they can even hinder effective management of personal information.

This research study thoroughly investigated the process of organizing personal information, including the actions people take, cognitive processes people go through, decisions people make, and factors that influence the process. Together, these results

make direct contributions to the development and design of various personal information strategies, devices, tools, interfaces, and applications that support individuals' organizing their information. When the findings of this study are incorporated into systems, they will result in more effective management of personal information to allow individuals to make better use of their resources, improving the quality and productivity of their lives.

7.3 Limitations

This research study was a case study which used a qualitative research method in exploring the process of organizing personal information. The qualitative research method was used because little was known about this area (i.e., the process of organizing personal information), so that it seemed appropriate and necessary to discover and explore a range of behaviors. It was an appropriate method in pursuing research objectives of this study, yet it has its limitations, in that the findings of this study cannot be said to be typical at this point. In addition, although the researcher tried to balance other variables as much as possible, there can be effects of other variables such as participants' gender, ethnicity, or age groups on the process of organizing personal information.

In addition, although a combination of the background questionnaire, a diary that was kept over a week, and two interviews provided rich data in examining the process of personal information organization, it was impossible to have a full description of the process of organizing all personal information items of each participant, as this study investigated the behavior at certain points of time. To be more specific, although this study examined the process of organizing each personal information that was recorded in

the diary during several days, and interviewed over weeks, what happened before or after those periods of time could not be fully examined.

Also, there is a possibility that this study itself influenced participants' behaviors, so that participants organized information items differently than usual. For instance, it is possible that participants might have tried to reveal themselves as more organized, or vice versa. To minimize the effect of the research setting and investigate participants' personal information organizing behavior in a natural setting, the researcher indicated in the informed consent form that the aspects of behavior that interested the researcher were how people organize their personal information, what decisions were made during the process, why certain decisions were made, what factors affected the decisions, and what people were thinking while organizing information, rather than evaluating participants' behavior against any standard measure such as good or bad organization. In addition, when participants were concerned about their unique organizing behavior, the researcher assured them that it was totally fine, and asked them to do things just the way they did in their daily lives.

Another limitation of this study was that it primarily relied on the verbal protocol (i.e., what was mentioned by participants) in exploring the process of organizing personal information. Thus, there is a possibility that sometimes participants failed to give a full, accurate, or clear description of the process. In fact, because moving from one stage to another happens very quickly, continuously, and often unconsciously, certain stages were indiscernible or not mentioned by the participants, even though it was evident that they went through this stage. For instance, although it was impossible to initiate any organization process without having or saving an information item, sometimes

participants did not mention that they saved the file or obtained the file in a certain way. Also, although it was evident that the participants clicked the folder to save an information item into the folder, the researcher could not identify those actions based on the verbal protocol. To elicit data about those behaviors, it seems that an eye tracking method or computer data logging method needs to be used.

7.4 Future Research

This research was an exploratory study that investigated the process of organizing personal information. Thus, the findings of this study can be further analyzed and extended to have a better understanding of the process of organizing personal information. This research study can be extended by examining and comparing the personal information organizing process of different social groups, such as academics in different disciplines or people in a particular industry, to further examine the social aspects of the process. This research study also can be extended by including different forms of personal information other than digital forms. For instance, how people who are ill organize their physical and digital medical information, and how physical and digital information organizations influence each other will be interesting to examine. In addition, the personal information organizing process of specific age groups, genders, or cultures will be interesting and practical areas to explore. Also, another research method such as eye tracking method or computer data logging method can be used as a supplemental method to further investigate the process.

There were various other aspects of information organization behavior that were found and coded while analyzing data. These aspects included different uses of various

personal devices, the influence of technology on information organization (such as the effect of using software that synchronizes categories across multiple personal information devices), usefulness of current interfaces and tools for digital information organization, and individual differences in their tendency to maintain their organizational structure. It will be worthwhile to analyze these data to further understand personal organizational behavior. Also, it will be interesting to analyze the characteristics as well as the patterns of organizing information items that are considered easy or difficult when making organizational decisions, or to explore the process of organizing information items with different uses.

In addition, it will be desirable to revisit participants to examine changes made to the files and folders described in the study, and to examine how personal information organizational structure develops over a longer period of time to further investigate the process, and strengthen the holistic view in exploring the process of organizing personal information.

APPENDICES

Appendix 1

Recruitment Letter

To: Students and Faculty members

From: Kyong Eun Oh

Subject: Participation in Personal Information Organization Process Study

Hello, my name is Kyong Eun Oh who is a doctoral student at School of Communication & Information at Rutgers University. For my dissertation, I am examining how academics in social science field organize their information in digital forms. I am looking for professors and students to participate in this study.

If you participate in this study, you will be asked to conduct three separate sessions.

- 1) A short interview, which asks brief background information. This interview will take about 5 minutes.
- 2) Recording a short diary entry for a week in a given template whenever you decide to save and/or organize an information item in a digital form.
- 3) Two interviews. The first interview will ask questions about how you saved and organized information in digital forms, and the second interview will ask whether there is any change in the files and folders since the last interview. The first interview will take about 40 minutes, and the second interview will take about 10 minutes.

If, at any time, information that you decided to keep is private or confidential, you may exclude it from the diary and the interview.

Specific date, time and place of the interview will be scheduled at your convenience.

The aspects of behavior that interest me are how academics organize their information in digital forms, what decisions are made during the process, why certain decisions are made, what factors affect the decisions and what people are thinking while organizing information. I am not interested in evaluating your behavior against any standard measure such as good or bad organization. Thus, there is no evaluation of individual performance. For all three sessions, your identity will remain strictly confidential.

The success of this dissertation relies upon academic participants, so your participation will be greatly appreciated. Your participation will help us to learn more about information organizing behaviors, and make contributions to the development and design of various personal information management devices, gadgets and interfaces that support an individual's organizing personal information.

If you are interested in participate in this study or want to know more about this study, please contact me at keoh@eden.rutgers.edu

Thank you very much.

Best wishes,
Kyong Eun Oh

Kyong Eun Oh
Doctoral Candidate
Department of Library and Information Science
School of Communication and Information
Rutgers University
4 Huntington Street
New Brunswick, NJ 08901-1071 USA

Appendix 2

Informed Consent Form

You are invited to participate in a research study that is being conducted by Kyong Eun Oh, who is a doctoral student in the School of Communication & Information at Rutgers University. The purpose of this research is to explore the process of organizing personal information in digital forms. By “personal information” is meant information that you keep for your own use either directly or indirectly. For instance, word files you received from someone via email attachment, image files that you saved while searching the Web, PDF files that you saved from electronic journals, excel files that you created for your own project, digital photos that you downloaded from digital camera or cell phone, music files that you ripped to your laptop from CD, etc.

The aspects of behavior that interest me are how people organize their personal information, what decisions are made during the process, why certain decisions are made, what factors affect the decisions and what people are thinking while organizing information. I am not interested in evaluating your behavior against any standard measure such as good or bad organization.

You must be 18 years or older and a faculty member or a graduate student or an undergraduate student in an institute of higher education in order to participate in this study. This study is recruiting approximately 20-24 participants.

Your participation will require three separate sessions with me, and an intervening week of your recording information about your information saving activities. At the first session, that is, today, I will explain to you more about your participation, ask brief background information about you, and explain how to record a diary over the next week. This will entail your making a diary entry whenever you decide to save an information item in a digital form that you encountered from various sources in any of your personal devices. You will also be asked to record a diary entry whenever you decided to put or move already saved information into a certain folder or a subfolder or application. If, at any time, information that you decided to keep is private or confidential, please exclude it from the diary. To record diary entries, you will be given a template which is in an excel file format. Thus, you will need to install an excel file on your computer.

The second session will be an interview that will take approximately half an hour. The first interview will be conducted shortly after you are done with recording the diary. In the interview, I will ask questions about how you kept information in your personal devices while looking at the diary that you wrote.

The third session will be another interview, which will be conducted two to four weeks after the first interview. In this interview, if there is any change in the files and folders that you kept and/or organized since the last interview and the reasons for the change will be asked.

Specific date, time and place of the interview will be scheduled at your convenience. If, at any time, we come across any private or confidential information that you do not want to include in the study, please exclude it. The interview will be recorded and transcribed if you agree to the recording of the interview.

At any time during this research, you may ask to withdraw from the study.

This research is anonymous. Anonymous means that I will record no information about you that could identify you. This means that I will not record your name, address, phone number, date of birth, etc. If you agree to take part in the study, you will be assigned a random code number that will be used on each test and the questionnaire. Your name will appear only on a list of subjects, and will not be linked to the code number that is assigned to you. There will be no way to link your responses back to you. Therefore, data collection is anonymous.

The research team and the Institutional Review Board at Rutgers University are the only parties that will be allowed to see the data, except as may be required by law. If a report of this study is published, or the results are presented at a professional conference, only group results will be stated. All study data will be kept for ten years and then destroyed.

You will benefit from our research in that by participating you will have greater understanding of your personal information organizing behavior, and make contributions to the development and design of various personal information management devices, gadgets and interfaces that support an individual's organizing personal information. However, you may receive no direct benefit from taking part in this study. There are no foreseeable risks to participation in this study higher than the general risk of using a computer.

Participation in this study is voluntary. You may choose not to participate, and you may withdraw at any time during the study procedures without any penalty to you. In addition, you may choose not to answer any questions with which you are not comfortable.

If you have any questions about the study or study procedures, you may contact me at

Kyong Eun Oh
School of Communication & Information
Department of Library and Information Science
Rutgers University
New Brunswick, NJ 08901
Tel: (732) 609-1106
Email: keoh@eden.rutgers.edu

Or you can contact my advisor Nicholas J. Belkin at

Nicholas J. Belkin
School of Communication & Information
Department of Library and Information Science
Rutgers University
New Brunswick, NJ 08901
Tel: 732-932-7500 ext. 8271
Email: belkin@rutgers.edu

If you have any questions about your rights as a research subject, you may contact the IRB Administrator at Rutgers University at:

Rutgers University, the State University of New Jersey
Institutional Review Board for the Protection of Human Subjects
Office of Research and Sponsored Programs
3 Rutgers Plaza
New Brunswick, NJ 08901-8559
Tel: 848-932-0150
Email: humansubjects@orsp.rutgers.edu

You will be given a copy of this consent form for your records.

Sign below if you agree to participate in this research study:

Subject (Print) _____

Subject Signature _____ Date _____

Principal Investigator Signature _____ Date _____

Appendix 3

Audio/Videotape Addendum to Consent Form

You have already agreed to participate in a research study entitled: Personal information organization process conducted by Kyong Eun Oh. We are asking for your permission to allow us to audiotape as part of that research study. You do not have to agree to be recorded in order to participate in the main part of the study.

The recording will be used for transcription and analysis. The recording will not include your name or any other identifier.

The recordings will be stored in a password protected folder on the principle investigator's computer with no link to subjects' identity, and will be destroyed upon publication of study results.

Your signature on this form grants the investigator named above permission to record you as described above during participation in the above-referenced study. The investigator will not use the recordings for any other reason than those stated in the consent form without your written permission.

Subject (Print) _____

Subject Signature _____ Date _____

Principal Investigator Signature _____ Date _____

Appendix 4

Initial Background Questionnaire

- Q1. What is your gender?
- Q2. What age group are you in?
- Q3. What is your ethnicity?
- Q4. What is your field of study?

Appendix 5

Diary Instruction

Thank you for agreeing to participate in this study. The purpose of this study is to explore the process of organizing personal information in digital forms.

By “**personal information**” I mean information that you keep for your own use either directly or indirectly. For instance, word files you received from someone via email attachment, image files that you saved while searching the Web, PDFs file that you saved from electronic journals, excel files that you created for your own project, digital photos that you downloaded from digital camera or cell phone, music files that you ripped to your laptop from CD, etc.

For one week, please set aside 1-2 minutes and record a diary entry in the following format whenever you decide to save an information item in a digital form.

Information item in digital form may include a Word file, text file, excel file, PowerPoint slides, music file, image file, PDFs, etc.

You may encounter them from various **sources** such as Email, Website, digital camera, CD, etc.

Then, you may save it into any of your personal **devices** such as laptop, desktop, smart phone, external drives, etc.

Please also record a diary entry whenever you decide to put or move already saved information into a certain folder or a subfolder or application.

Please do as you usually do when saving or organizing information that you keep for your own use.

If, at any time, information that you decided to keep is private or confidential, please exclude it from the diary.

Please record incidents completely while using as much detail as possible.

Thank you very much for your participation.

[illegible]

Appendix 7

Initial First Interview Questions

[For each entry in the diary]

- Q1. Why did you decide to save this information item in that device?
- Q2. According to what you did while saving the information item,
- Q2-1. If you decided to save it in the one of the existing folders,
 Why did you save it in that folder?
 What other things are in that folder?
 - Q2-2. If you decided to create a new folder for the information item,
 Why did you create a new folder?
 - Q2-3. If you moved it to another folder,
 Why did you move it to another folder?
 What other things were in the previous folder?
 What other things are in the folder that you moved information item?
 - Q2-4. If you changed the name of the folder,
 Why did you change the name of the folder?
 What other things are in that folder?
 - Q2-5. If you did not put it into a specific folder,
 Why did you put it there?
 What other things are there?
- Q3. What is the main use of this information item? (What is this information item for?)
- Q4. Was it easy for you to decide where to save this information item?
 If so, why was it easy? If not why was it not easy?
- Q5. Can you think about any other existing folders you can save this information item?
 If so, why did you think this folder (location) is better than that folder in saving this information item?
- Q6. Do you think there is a possibility of moving this information item into other place sometime later? If so, why do you think so? If not, why do you think so?

Appendix 8

Second Interview Questions

[For each entry in the diary]

Q1. Did you made any change to this information item since the last interview?

Q1-1. If there was any change, what did you do?

Q1-1-1. Why did you make that decision?

Appendix 9

Revised Background Questionnaire

- Q1. What is your gender?
- Q2. What age group are you in?
- Q3. What is your ethnicity?
- Q4. Are you an undergraduate student? Graduate student? Post-doctoral researcher? Or professor?
- Q5. How many years have you spent at Universities?
- Q6. What is your field of study?
- Q7. Other than being a student or a professor, what other primary roles do you have in your daily life?
- Q8. Would you say your job as an academic is a very important source of identity for you? Why?
- Q9. Among your primary roles, which role do you regard the most important source of identity for you? Why?

Appendix 10

Revised First Interview Questions

Q1. In your daily life, in which devices do you usually keep your information items? Ex. laptop or desktop at home or school, external drive, USB, mobile phone, iPad, etc.

[For each entry in the diary]

Q2. Why did you decide to save this information item in that device?

Q3. According to what you did while saving the information item,

Q3-1. If you decided to save it in the one of the existing folders,

Why did you save it in that folder?

What other things are in that folder?

When and why did you create this folder?

Q3-2. If you decided to create a new folder for the information item,

Why did you create a new folder?

Q3-3. If you moved it to another folder,

Why did you move it to another folder?

What other things were in the previous folder?

What other things are in the folder to which you moved the information item?

Q3-4. If you changed the name of the folder,

Why did you change the name of the folder?

What other things are in that folder?

Q3-5. If you did not put it into a specific folder,

Why did you put it there?

What other things are there?

Q4. What is the main use of this information item? (What is this information item for?)

Q4-1. If for academic work, in what role?

Q4-2. If for other work, in what role?

Q4-3. If for personal use, was it for yourself, or to share with others?

Q5. Was it easy for you to decide where to save this information item?

If so, why was it easy? If not why was it not easy?

Q6. Suppose you are not allowed to save this information item in this folder (location).

Where would you save this information item, instead?

Why is this folder (location) better than other folders (locations) in saving this information item?

Q7. Do you think there is a possibility of moving this information item into other place sometime later? If so, why do you think so? If not, why do you think so?

Appendix 11

Part of the Coding Scheme

P #	Group	Subgroup	Info ID	File Name	Thought Community	Device	Format	Stage	Code level 1	Code level 2	Code level 3	Line #	Quote
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	1	Action	create file		325-326	So that's the naming conventio
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	2	Action	identify file		320	So that's a tutorial. That's for a
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	5	Action	select existing category		331	I saved it under teaching.
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	6	Action	place a file into a category		331	I saved it under teaching.
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	1	Decision	organize		328	And this time, you saved it in T
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	5	Decision	select		331	I saved it under teaching.
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	5	Decision	select		331-334	So under teaching, once again,
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	5	Decision	select		365-366	So this tutorial will not be unde
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	6	Decision	keep categorization		372	I'm happy with where it is.
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	4	Factor	purpose		330-331	So it's a conference thing, but i
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	4	Factor	purpose		355-356	As I said, it's not really a paper
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	4	Factor	purpose		356-363	And one way to kind of think al
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	2	Factor	purpose		320	So that's a tutorial. That's for a
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	6	Factor	satisfaction		372	I'm happy with where it is.
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	5	Factor	purpose		331-334	So under teaching, once again,
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	5	Factor	purpose		365-366	So this tutorial will not be unde
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	2	Factor	purpose		320	So that's a tutorial. That's for a
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	2	Thought	typification		320	So that's a tutorial. That's for a
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	4	Thought	Assess similarities and differences bet		330-331	So it's a conference thing, but i
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	4	Thought	Assess similarities and differences bet		355-356	As I said, it's not really a paper
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	4	Thought	Assess similarities and differences bet		356-363	And one way to kind of think al
P22	Professor	pretenure	2203	tut07-P22	Academia	Laptop	text	5	Thought	Adjust mental gap between new and e		331-334	So under teaching, once again,

Appendix 12

Codebook

Personal Information Organization Process Model

1. Please identify an information item in the transcript.
2. Please identify a category where an information item is saved.
3. Code <Code Level 1> by identifying any Action, Thought, Decision, or Factor that is involved in the process of organizing an information item.
4. Then, code <Code Level 2> for the specification by referring to the table in this code book.
5. If there is additional explanation you want to code, code it in <Code Level 3>
6. Identify which <Stage> each code belongs except for the 'FACTORS'.

<Code Level 1>

ACTION
THOUGHT
DECISION
FACTORS

<Code Level 2 >

See the table

<Stage>

1 INITIATION
2 IDENTIFICATION
3 TEMPORARY CATEGORIZATION
4 EXAMINATION/COMPARISON
5 SELECTION/MODIFICATION/CREATION
6 CATEGORIZATION

CODE LEVEL 1: ACTION

STAGE	CODE LEVEL 2	DESCRIPTION	EXAMPLE
1	Receive file	The participant talks about receiving or getting file(s) from other person.	"I also received these files." "Finally got all files from that people by email." "My friend emailed them to me."
	Create file	The participant talks about creating new file(s).	"I created this file." "It's the notes that I took during my lecture."

			<p>“Interviewer: So you created it while taking a class?”</p> <p>Interviewee: Yeah.”</p>
	Save file	<p>The participant talks about saving file. This includes downloading file from email or web site, uploading file, bookmarking a web site, or taking a screen shot. This can be overlapped with one of the quote(s) for DECISION in Stage 1 which is ‘save file’. This can be overlapped with the quote(s) for ACTION in Stage 3, which is ‘save in a temporary location’. This also can be overlapped with the quote(s) for one of the ACTION or DECISION in Stage 5 which is ‘select a folder’.</p>	<p>“I saved it in Conference Name (folder).”</p> <p>“I just saved it to my Desktop.”</p> <p>“I downloaded them.”</p>
	Obtain file in another way	<p>The participant talks about other ways of having file(s) such as by restoring file(s) from a trash can, or copying existing file(s).</p>	<p>“But I um... restored it about a few days ago.”</p> <p>“I took it from a file (folder) called ‘Teaching Course 1’, and I saved it to Desktop.”</p>
3	Save file in a temporary location	<p>The participant mentions saving files in a certain location, usually easily accessible location, for a short period of time.</p>	<p>“I put it on the ‘Desktop’.”</p> <p>“I just temporarily put it on the ‘Desktop’.”</p> <p>“I just left it in ‘Downloads’.”</p> <p>“When you download something, it automatically put it in ‘Download’ folder.”</p>
5	Select a folder	<p>The participant mentions saving file into one of the existing folder. Also, based on the name of the folder(s), please code ‘factor(s)’ that impacted on the action. This can be overlapped with the quote(s) for DECISION in Stage 1 which is ‘save file’. This</p>	<p>“I saved it under ‘Teaching’.”</p> <p>“I saved in ‘Research’ folder.”</p> <p>“I put it in my ‘Exam 2’ folder.”</p>

		also can be overlapped with the quote(s) for one of the DECISION in Stage 5 which is 'select a folder'.	
	Modify previous folder	The participant mentions splitting a folder (or creating a subfolder), merging folder (or creating a super folder), or changing the name of the folder.	<p>"So, I took those files, and I split them up into the ones I was assigned to review." (split)</p> <p>"I made sub folders." (split)</p> <p>"I actually created a new folder called Travel picture folder and then I put it there. I put the folder into the bigger category." (merge)</p> <p>"I changed the name of the folder." (change the name)</p>
	Create a new folder	The participant mentions creating a new folder to save file.	<p>"I created a new folder there."</p> <p>"I created this 'Colleague name' folder."</p> <p>"Interviewer: At this time you saved it in a new folder called 'Vacation Info'.? Interviewee: Yes."</p>
6	Place file into a folder	Participants talks about placing a file into a folder. This often can be overlapped with the quote(s) for 'ACTION' and 'DECISION' in Stage 5 which is 'selecting a folder'.	<p>"So I saved the excel file of their grades under the same folder, name 'Exams'"</p> <p>"I saved it in my 'Spring Semester' folder."</p> <p>"It goes in the 'Dissertation' folder."</p>

CODE LEVEL 1: THOUGHT

STAGE	CODE LEVEL 2	DESCRIPTION	EXAMPLE
1	Messy	The participant talks about feeling messy or unclear when he/she sees file(s).	<p>"Things got so messy, that I just threw them all in this folder."</p> <p>"It's pretty big and messy."</p>
	Confusing	The participant talks about feeling confused when he/she sees file(s).	"I realized lot of my stuff was very confusing and everything was like all over the place in the 'Documents'?"
2	Typify	The participant typify file(s) as "something" while	"That's my dissertation survey."

		focusing on certain aspect(s) of the file such as format, purpose, related person, source, time, topic, type, or use of that file(s).	<p>“It’s a PDF.”</p> <p>“It is a review.”</p> <p>“It was for class.”</p>
4	Review existing folders	The participant mentions any of his/her existing folders or folder structure that are relevant to unorganized file.	<p>“That’s the folder I organize my summer teaching materials.”</p> <p>“I keep all my survey, like survey instruments and revised surveys under ‘Dissertation/IRB folder’.”</p> <p>“I’ve organized it by different research projects that I’m working on.”</p>
	Assess similarities and differences between new and existing files	The participant mentions similarities or differences between unorganized file and existing files	<p>“They are notes, I don’t want notes with my homework assignments.”</p> <p>“I saved all of the papers in that folder so the directions for how to review those papers should probably be saved at that folder, too.”</p>
5	Adjust mental gap between new and existing files	The participant mentions 1) selecting or 2) deselecting one of the existing folders, 3) modifying existing folder(s) or 4) creating a new folder by describing either similarities or differences between unorganized and organized files in relevant folder(s).	<p>“Because it was for that class, um, it was an assignment for the class, so it would go in the class folder.”</p> <p>“So I just put it all in the same folder because, it was regarding the same information.”</p> <p>“I don’t know they are notes, I don’t want notes with my homework assignments.”</p> <p>“It’s a totally new job, so I don’t want to conflict with my existing folders.”</p>
6	Clean	The participant says it is clean, neat, simple, or not messy when files are organized into folders. Or participant uses the expression ‘clean up’ to describe organizing files.	<p>“I may take whole day and to clean up on it, sometimes it gets kind of messy.”</p> <p>“That folder needs a bit of clean up. I think I should put the talks in by year. ‘Cause it’s pretty big and messy.”</p> <p>“When I clean it, I can just put it into the right folder.”</p>

CODE LEVEL 1: DECISION

STAGE	CODE LEVEL 2	DESCRIPTION	EXAMPLE
1	Save file	The participant talks about saving file(s) into folder(s) or organizing file(s) in a certain way. This can be overlapped with one of the quote(s) for ACTION in Stage 1 which is 'save file'. This can be overlapped with the quote(s) for ACTION in Stage 3, which is 'save in a temporary location'. This also can be overlapped with the quote(s) for one of the ACTION or DECISION in Stage 5 which is 'select a folder'.	"I saved them in different folders." "I just saved it to my 'Desktop'." "Interviewer: You saved it in a laptop, 'Documents'? Interviewee: Yes."
3	Delay decision making	The participant talks about not making any organizational decision for the time being, or mentions future decision he/she is going to make. This often involves expressions 'will' or 'future'. However, if the participant mentions possibility usually with expressions such as 'maybe' or 'can', it should not be coded as 'Delay decision making'.	"I keep it in my Download folder for the time being." "And all of these files will go in that folder. I just haven't filed them yet." "I'll put it in a folder for the Summer session 2012." "I'll actually probably just delete it."
	Re-categorize file	In the second interview, the participant mentions that he/she moved the file(s) that was(were) saved in a temporary location to a folder.	"I moved it into the folder here." "I moved it into my 'Spring Semester' folder."
	Delete file	In the second interview, the participant mentions that he/she deleted the file(s) that was(were) saved in a temporary location.	"I threw them away." "I deleted this." "And then I erased it."

	Keep temporary categorization	In the second interview, the participant mentions that the file(s) that was(were) saved in a temporary location is still in that location, or he/she will keep it there for a while.	<p>“It’s still there.”</p> <p>“It’s still on my Desktop. And I haven’t done the work.”</p>
5	Select a folder	<p>The participant mentions saving file into one of the existing folders.</p> <p>Even if the participant only talks about saving file without mentioning specific folder, if it was saved into a folder, the quote needs to be coded. In this case, the path name of the folder is in parenthesis.</p> <p>Also, based on the name of the folder(s), please code ‘factor(s)’ that impacted on the decision.</p> <p>This can be overlapped with the quote(s) for ACTION or DECISION in Stage 1 which are ‘save file’, ‘receive file’, or ‘obtain file in another way’. This also can be overlapped with the quote(s) for one of the ACTION in Stage 5 which is ‘select a folder’.</p>	<p>“I put it under ‘Reviews’.”</p> <p>“I saved them under ‘Travel’.”</p> <p>“I filed it under ‘Fall Courses’.”</p> <p>“Interviewer: And you saved it in ‘Conference 2012’ folder? Interviewee: Yeah.”</p> <p>“So I saved it in this file (Documents\SpringLake).”</p>
	Modify previous folder	<p>The participant talks about modifying (splitting, merging, creating a subfolder, or change the name of the folder) one of the existing folders to save file.</p> <p>If the participant mentions that he/she re-categorized the folder in the second interview, it needs to be coded as well.</p>	<p>“I made subfolders.”</p> <p>“I saved it in a folder that I already had, but you can see I changed the name of the folder.”</p> <p>“I did make a week 2, 3, and 4 and moved all those stuff that are downloaded for those into them.”</p>
	Create a new folder	The participant mentions creating a new folder to save	“I created this ‘Final analysis’ folder.”

		file.	“And then, I created a folder there.”
6	Delay decision making	In the second interview, the participant talks about delaying making organizational decision to the file(s) that was(were) saved in a folder, or mentions future decision he/she is going to make. This often involves expressions ‘will’ or ‘future’. However, if the participant mentions possibility usually with expressions such as ‘maybe’ or ‘can’, it should not be coded as ‘Delay decision making’.	“I actually at the end of the semester will just combine all my folders because I don’t need them for classes anymore.” “it’s going to get moved.” “I will delete it when I’m done with everything.”
	Re-categorize file	In the second interview, the participant talks about re-categorizing file(s) that was(were) saved in a folder by 1) moving it to different folder, 2) creating a subfolder (split), or 3) merging into a bigger folder (merge). If it was split or merged, it also need to be coded as ‘modify previous folder’ in both ACTION and DECION for Stage 5.	“That’s another one of those letters that got filed into ‘Travel’.” “I put the ones that have to do with one class into one folder. So I did that for a few of my classes.” “I put it under the folder ‘Quiz #2’, again.” “I made another folder for it.”
	Delete file	In the second interview, the participants mentions that he/she deleted the file(s) that was(were) saved in a folder.	“I actually deleted the file.” “I erased the whole folder, like, a couple days ago actually.”
	Keep categorization	In the second interview, the participant talks about keeping organized file into the folder without making any changes.	“That one is still in the ‘Documents’ folder.” “That’s so same, I didn’t move.” “It’s still just left in that folder.”
	Move file	In the second interview, the participant talks about keeping organized file into the folder but moving the	“I decided to just copy this video from my laptop to this hard drive.” “Interviewer: So you moved

		folder into another personal device.	it to a shared drive? Interviewee: Yes.”
--	--	--------------------------------------	---

CODE LEVEL 1: FACTORS

CODE LEVEL 2	DESCRIPTION	EXAMPLE
Accessibility	The participant mentions the ability to access, find, or browse file(s) that impacted on any stage of the process of organizing personal information.	<p>“It’ll be easy to find it.”</p> <p>“A lot of it is just keeping things accessible and relatively easy to find.”</p> <p>“I just saved it to the easiest place that I can locate.”</p> <p>“I created it so that I could put all my stuffs in that class in this folder, so it would be, you know, easy to find. I would know where to find it.”</p>
Affiliation	The participant mentions the impact of any group or institution in which participants are formally connected or joined that is relevant to the file(s) on any stage of the process of organizing personal information. This includes the particular name of university, school, department, or program. However, if the participant only said ‘school’ such as ‘school thing’, it needs to be coded as ‘purpose’. In addition, if the participant mentions about year in school such as ‘freshman’ or ‘junior’, it needs to be coded as ‘time’.	<p>“It’s under ‘University 1’ (folder) because it’s my first time to work as a translator here.”</p>
Appropriateness	The participant mentions the appropriateness or inappropriateness that impacted on any stage of the process of organizing personal information. This also includes satisfaction or dissatisfaction about organization.	<p>“It’s really related to exactly that. It’s really specific to that reviewing assignment. So that’s the only place I will store it.”</p> <p>“I’m happy with where it is.”</p>
Availability	The participant mentions the availability or unavailability of categories which impacted on any stage of the process of organizing personal information.	<p>“I didn’t have a folder yet for that company.”</p> <p>“That’s because I don’t really have anywhere else.”</p>

Format	The participant talks about the physical characteristics including digital format or medium of file(s) that impacted on any stage of the process of organizing personal information.	<p>“It’s a Word file.”</p> <p>“It’s a PDF.”</p> <p>“I save all my photos in iPhoto.”</p> <p>“Because it’s a picture. It’s a Picture folder. It’s a folder for pictures.”</p>
Messiness	The participant talks about messiness or confusion that impacted on any stage of the process of organizing personal information.	<p>“I did it partly because there was too much mess on my screen.”</p> <p>“It’s maybe four or five files. So I thought that’s enough that it doesn’t make a mess.”</p> <p>“I realized lot of my stuff was very confusing.”</p>
Necessity of differentiation	The participant talks about the need for distinguishing certain files from other files that impacted on any stage of the process of organizing personal information.	<p>“I put it in the ‘exam 2’ folder, because that was what I needed to know for exam 2, and I don’t want to have it mixed up with the exam 3 notes that I need to study.”</p> <p>“It’s a totally new job, so I don’t want to conflict with my existing folders.”</p>
Number of files	The participant talks about few or many numbers of related files which impacted on any stage of the process of organizing personal information. This also includes anticipation or no anticipation of having more files.	<p>“I had so many files for that specific class. I cannot like spread all the files in the ‘Desktop’.”</p> <p>“I just left it in ‘Downloads’ because I haven’t created a folder for this class yet because I haven’t had much for it.”</p> <p>“Just because I figured there would be more things that would be going into that folder.”</p>
Purpose	The participant talks about the purpose of file(s), or task that is related to the file which impacted on any stage of the process of organizing personal information. If the participant mentions that the file(s) was(were) ‘used for something’, it needs to be coded. Usually, participants use the expression ‘to’ or ‘for’ in	<p>“It’s for a Writing class.”</p> <p>“There’s also papers that are associated with papers that he is writing.”</p> <p>“It’s an assignment.”</p> <p>“So it’s a conference thing, but it’s not really a publication or a paper.”</p> <p>“It’s a totally new job.”</p> <p>“It’s for my Dissertation. So it</p>

	<p>describing the purpose of the file(s). ‘class’, ‘project’, ‘paper’, ‘review’, ‘exam’ need to be coded as purpose. If the participant mentions the name of the class, paper, or project, it also needs to be coded as ‘topic’.</p>	<p>goes in the ‘Dissertation’ folder.”</p>
Related person	<p>The participant talks about the impact of anyone who is related to file(s) on any stage of the process of organizing personal information. This includes 1) when they worked on an information item with someone, 2) when it is someone’s work, 3) when the contents of an information item is related with someone, and 4) when an information item is for someone or 5) from someone.</p>	<p>“That’s Daniel’s paper.” “It’s for my advisor.” “There is folder for graduate students, graduate students I’ve had. And so he would be inside of that folder”</p>
Reminder	<p>The participant talks about saving file(s) in a certain place as a visual reminder.</p>	<p>“I saved it on the ‘Desktop’ to remind me that I got to work on the revisions for her paper.” “I did it as a visual reminder.” “I put it on, in a folder on my Desktop to remind me I have papers to give him feedback on it.”</p>
Source	<p>The participant talks about the source of file(s) - where the file(s) came from - that impacted on any stage of the process of organizing personal information. The source of file refers to where participant obtained the file. Usually, participants use the expression ‘from’ to describe the source of the file(s).</p>	<p>“That was some other stuff that Conference A sent me.” “This was also a PDF file from the Web.” “I got the file from Google Scholar.” “They are all from my father.”</p>
System	<p>The participant talks about the impact of automated system or the default setting of a software, web browser, or personal device that impacted on any stage of the process of organizing personal information.</p>	<p>“It automatically goes into ‘Downloads’ folder.” “It automatically makes the folder with the artist’s name on it.” “I plug it and it automatically uploads anything new.”</p>

Time	The participant talks about the time dimension that is related to the file(s) that impacted on any stage of the process of organizing personal information. This includes certain point of time (year, semester, season, month, week, or day) that is relevant to the file as well as recentness (last, new, old) of the file(s).	<p>“This is the demo from 2011.”</p> <p>“So I basically organize it by when I’m in school or not in school.”</p> <p>“I have folder for like, Freshman Year, sophomore and High School folders.”</p> <p>“I keep all of my spring semester classes in one place.”</p> <p>“This is the paper that I worked on last year, last summer.”</p>
Time availability	The participant talks about having or not having enough time during the process of organizing personal information. This also includes saving time that impacted on the process.	<p>“At some point, I just ran out of time, I had too much to do.”</p> <p>“I just was in a rush when I was working on it, so I just saved it quickly to my ‘Desktop’.”</p> <p>“A lot of it is just keeping things accessible and relatively easy to find without taking the time.”</p>
Topic	The participant talks about the topic or subject of file(s) that impacted on any stage of the process of organizing personal information. If the participant mentions that the file(s) is(are) ‘about something’, it needs to be coded. This includes the specific name of the project or class as well as the general contents of the file.	<p>“I downloaded this which is an article about peer friendships.”</p> <p>“And the paper or the poster was on the CIS topic.”</p> <p>“It’s my office hours.”</p> <p>“This is the rules for this trip.”</p>
Type	The participant talks about the general type or genre of file(s) which impacted on any stage of the process of organizing personal information. ‘assignment’, ‘paper’, ‘note’ needs to be coded as ‘purpose’.	<p>“It’s an article.”</p> <p>“This is the actual template.”</p>
Use	The participant mentions about the use of file(s) that impacted on any stage of the process of organizing personal information. This includes (no) ongoing use, (no) future use, (no) possible use, and short or long-term use.	<p>“You asked why did I save it? For future use.”</p> <p>“Because I’m still working on it.”</p> <p>“Because I need to print it out.”</p> <p>“So that’s also done.”</p>

		“Generally the reason that I save things in my ‘Desktop’ is that because I work on things, instantly.”
Value	The participant talks about the personal meaning of the file to the participant such as interesting, important, favorite, meaningless that impacted on the process of personal information organization.	<p>“It was just a funny thing that was happening where Google was reporting that Google.com not found. And so I just took a screen shot of it.”</p> <p>“They are not that important to save in my, like permanent folder.”</p>

BIBLIOGRAPHY

- Abrams, D., Baecker, R., & Chignell, M. (1998). Information archiving with bookmarks: personal web space construction and organization. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, USA*, 41-48.
- Aristotle. (2007). *Categories* (D.M. Edghill, Trans.). Adelaide, South Australia: eBooks@Adelaide, University of Adelaide Library. Available: <http://etext.library.adelaide.edu.au/a/aristotle/categories/> (Original work written B.C. 350).
- Bälter, O. (1997). Strategies for organizing email messages. In H. Thimbleby, B. O'Connell, & P. Thomas (Eds.), *Proceedings of the Conference of the British Computer Society Human Computer Interaction Specialist Group - People and Computers XII: Vol. 12*. (pp. 21-38). Bristol, UK: Springer.
- Bälter, O. (1998). *Electronic mail in a working context*. Unpublished doctoral dissertation, Royal Institute of Technology, Stockholm, Sweden.
- Barreau, D. (1995). Context as a factor in personal information management systems. *Journal of the American Society for Information Science*, 46(5), 327-339.
- Barreau, D. (2008). The persistence of behavior and form in the organization of personal information. *Journal of the American Society for Information Science and Technology*, 59(2), 307-317.
- Barreau, D., Capra, R., Dumais, S., Jones, W., & Perez-Quinones, M. (2008). Introduction to keeping, refinding and sharing personal information. *ACM Transactions on Information Systems*, 26(4), 1-3.
- Barreau, D., & Nardi, B. (1995). Finding and reminding: File organization from the desktop. *ACM SIGCHI Bulletin*, 27(3), 39-43.
- Barsalou, L. (1983). Ad hoc categories. *Memory and Cognition*, 11(3), 211-227.
- Bates, M. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5), 407-424.
- Belkin, N. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, 5, 133-143.
- Bellotti, V., Ducheneaut, N., Howard, M., Smith, I., & Grinter, R. (2005). Quality vs. quantity: Email-centric task-management and its relationship with overload. *Human Computer Interaction*, 20, 89-138.

- Berger, P., & Luckmann, T. (1966). *The social construction of reality: A treatise on the sociology of knowledge*. Garden City, NY: Doubleday & Company.
- Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press.
- Bergman, O., Beyth-Marom, R., Nachmias, R., Gradovitch, N., & Whittaker, S. (2008). Improved search engines and navigation preference in personal information management. *ACM Transactions on Information Systems*, 26(4), 1-24.
- Bergman, O., Whittaker, S., Sanderson, M., Nachmias, R., & Ramamoorthy, A. (2010). The effect of folder structure on personal file navigation. *Journal of the American Society for Information Science and Technology*, 61(12), 2426-2441.
- Boardman, R., & Sasse, M. (2004). "Stuff goes into the computer and doesn't come out" A cross-tool study of personal information management. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Austria*, 6, 583-590.
- Boardman, R., Spence, R., & Sasse, M. (2003). *Too many hierarchies?: The daily struggle for control of the workspace*. Paper presented at the HCI International 2003: International Conference on Human-Computer Interaction, Crete, Greece.
- Bondarenko, O., Janssen, R., & Driessen, S. (2010). Requirements for the design of a personal document-management system. *Journal of the American Society for Information Science and Technology*, 61(3), 468-482.
- Brekhus, W. (2007). The Rutgers school: A Zerubavelian culturalist cognitive sociology. *European Journal of Social Theory*, 10, 448-464.
- Bruce, H. (2005). Personal, anticipated information need. *Information Research*, 10(3), Available at <http://informationr.net/ir/10-3/paper232.html>.
- Bruce, H., Jones, W., & Dumais, S. (2004). *Keeping and re-finding information on the web: What do people do and what do they need?* Paper presented at the ASIST 2004: Annual Meeting of the American Society for Information Science and Technology, Providence, RI.
- Bush, V. (1945). As We May Think. *The Atlantic Monthly*, 176, 641-649.
- Case, D. (1986). Collection and organization of written information by social scientists and humanists: a review and exploratory study. *Journal of Information Science*, 12(3), 97-104.
- Case, D. (1991). Conceptual organization and retrieval of text by historians: The role of memory and metaphor. *Journal of the American Society for Information Science*, 42(9), 657-668.

- Chen, H., & Dhar, V. (1991). Cognitive process as a basis for intelligent retrieval systems design. *Information Processing & Management*, 27(5), 405-432.
- Cole, I. (1982). Human aspects of office filing: implications for the electronic office. *Proceedings of the Human Factors Society, USA*, 59-63.
- Cole, C. (1997). Information as process: The difference between corroborating evidence and 'information' in humanistic research domains. *Information Processing & Management*, 33(1), 55-67.
- Cooper, L. (2004). The socialization of information behavior: a case study of cognitive categories for library information. *Library Quarterly*, 74(3), 299-336.
- Cutrell, E., Dumais, S., & Teevan, J. (2006). Searching to eliminate personal information management. *Communications of the ACM*, 49(1), 58-64.
- Dabbish, L. A., Kraut, R., Fussell, S., & Kiesler, S. (2005). Understanding email use: Predicting action of a message. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, USA*, 691-700.
- DiMaggio, P. (1997). Culture and cognition. *Annual Review of Sociology*, 23, 263-287.
- Douglas, M. (1978). *Purity and danger: An analysis of concepts of pollution and taboo*. London, UK: Routledge & Kegan Paul.
- Ducheneaut, N., & Bellotti, V. (2001). E-mail as habitat. *Interactions*, 8(5), 30-38.
- Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., & Robbins, D. (2003). Stuff I've seen: A system for personal information retrieval and re-use. *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval, Canada*, 72-79.
- Ekman, P., Friesen, W., & Ellsworth, P. (1972). *Emotion in the Human Face*. Elmsford, NY: Pergamon Press.
- Ellis, D. (1989). A behavioral approach to information retrieval system design. *Journal of Documentation*, 45, 171-212.
- Elswailer, D., & Ruthven, I. (2007). Towards task-based personal information management evaluations. *Proceedings of the ACM SIGIR Conference, Netherlands*, 23-30.
- Erickson, M. (2005). *Science, culture and society: Understanding science in the 21st century*. Cambridge, UK: Polity Press.
- Erickson, T. (2006). From PIM to GIM: Personal information management in group contexts. *Communications of the ACM*, 29(1), 74-75.

- Fisher, D., Bruxh, A., Gleave, E., & Smith, V. (2006). Revisiting Whittaker & Sidner's "Email overload" ten years later. *Proceedings of the ACM CSCW Conference, Canada*, 309-312.
- Fleck, L. (1981). *Genesis and development of a scientific fact*. Chicago, IL: University of Chicago Press.
- Gonçalves, D., & Jorge, J. (2003). An empirical study of personal document spaces. *Proceedings of the International Workshop on Design Specification, and Verification of Interactive Systems*, 46-60.
- Gottlieb, L., & Dilevko, J. (2003). Investigating how individuals conceptually and physically structure file folders for electronic bookmarks: The example of the financial services industry. *Journal of the American Society for Information Science and Technology*, 54(2), 124-139.
- Gwizdka, J. (2004). Email task management styles: The cleaners and the keepers. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Austria*, 1235-1238.
- Henderson, S. (2005). Genre, task, topic and time: facets of personal digital document management. *Proceedings of ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction, New Zealand*, 94, 75-82.
- Henderson, S. (2009). *How do people manage their documents? An empirical investigation into personal document management practices among knowledge workers*. Unpublished doctoral dissertation, University of Auckland, Auckland, New Zealand.
- Henderson, S., & Srinivasan, A. (2009). An empirical analysis of personal digital document structures. *Human Interface and the Management of Information*, 56(17), 394-403.
- Iyer, H. (1995). *Classificatory structures: Concepts, relations, and representation*. Frankfurt, DE: Index Verlag.
- Jones, W. (2004). Finders, keepers? The present and future perfect in support of personal information management. *First Monday*, 9(3), Available at http://www.firstmonday.dk/issues/issue9_3/jones/index.html.
- Jones, W. (2007). Personal Information Management. *Annual Review of Information Science and Technology*, 41, 453-504.
- Jones, W. (2008). *Keeping found things found: The study and practice of personal information management*. San Francisco, CA: Morgan Kaufmann Publishers.
- Jones, W., Bruce, H., & Dumais, S. (2001). Keeping found things found on the web. *Proceedings of International Conference on Information and Knowledge Management, USA*, 119-126.

- Jones, W., Dumais, S., & Bruce, H. (2002). Once found, what then? : A study of "keeping" behaviors in the personal use of web information. *Proceedings of the Annual Meeting of the American Society for Information Science and Technology, USA*, 391-402.
- Jones, W., Phuwanartnurak, A. J., Gill, R., & Bruce, H. (2005, April). *Don't take my folders away! Organizing personal information to get things done*. Paper presented at the ACM SIGCHI Conference on Human Factors in Computing Systems, Portland, OR.
- Jones, W., & Teevan, J. (2007). *Personal information management*. Seattle, WA: University of Washington Press.
- Krathwohl, D. R. (1997). *Method of educational and social science research: An integrated approach*. New York, NY: Longman.
- Kuhlthau, C. (1983). *The library research process: Case studies and Interventions with high school seniors in advanced placement English classes using Kelly's theory of constructs*. Unpublished doctoral dissertation, Rutgers, The State University of New Jersey, New Brunswick.
- Kuhlthau, C. (1993). A principle of uncertainty for information seeking. *Journal of Documentation*, 49(4), 339-355.
- Kuhlthau, C. (2005). Kuhlthau's Information search process. In K. Fisher, S. Erdelez, & L. McKechnie (Ed.), *Theories of information behavior* (pp. 230-234). Medham, NJ: Information Today.
- Kwasnik, B. (1989). *The influence of context on classificatory behavior*. Unpublished doctoral dissertation, Rutgers, The State University of New Jersey, New Brunswick.
- Kwasnik, B. (1991). The importance of factors that are not document attributes in the organization of personal documents. *Journal of Documentation*, 47(4), 389-398.
- Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago, IL: The University of Chicago Press.
- Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics*, 19(1), 55-66.
- Lounsbury, F. (1956). A semantic analysis of the Pawnee kinship usage. *Linguistic Society of America*, 32(1), 158-194.
- Mackay, W. (1988). Diversity in the use of electronic mail: A preliminary inquiry. *ACM Transactions on Office Information Systems*, 6(4), 380-397.
- Malone, T. (1983). How do people organize their desks: implications for the design of office information-systems. *ACM Transactions on Office Information Systems*, 1(1), 99-112.

- Marshall, C., & Jones, W. (2006). Keeping Encountered Information. *Communications of the ACM*, 49(1), 66-67.
- Medin, D. L., & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review*, 85(3), 207-238.
- Nardi, B., Anderson, K., & Erickson, T. (1995, July 4-7). *Filing and finding computer files*. Paper presented at the International East-West International Conference on Human-Computer Interaction, Moscow, Russia.
- Naumer, C., & Fisher, K. (2007). Naturalistic approaches for understanding PIM. In W. Jones & J. Teevan (Ed.), *Personal information management* (pp. 76-88). Seattle, WA: University of Washington Press.
- Ratneshwar, S., Barsalou, L, Pechmann, D., & Moore, M. (2001). Goal-derived categories: The role of personal and situational goals in category representations. *Journal of Consumer Psychology*, 10(3), 147-157.
- Ravasio, P., Schär, S. G., & Krueger, H. (2004). In pursuit of desktop evolution: User problems and practices with modern desktop systems. *ACM Transactions on Computer-Human Interaction*, 11(2), 156-180.
- Rieh, S. (2000). *The information quality and cognitive authority in the World Wide Web*. Unpublished doctoral dissertation, Rutgers, The State University of New Jersey, New Brunswick.
- Rodden, K. (1999). *How do people organize their photographs?* Paper presented at the Annual Colloquium on Information Retrieval Research, Glasgow, Scotland.
- Rodden, K., & Wood, K. (2003). *How do people manage their digital photographs?* Paper presented at the SIGCHI Conference on Human factors in computing systems, Ft. Lauderdale, FL.
- Rosch, E. (1978). Principles of categorization. In E. Rosch, & B. Lloyd (Ed.), *Cognition and Categorization* (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories, *Cognitive Psychology*, 7, 573-605.
- Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories, *Cognitive Psychology*, 8, 382-439.
- Rosch, E., Simpson, C., & Miller, R. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 491-502.

- Russell, D., & Lawrence, S. (2007). Search everything. In W. Jones & J. Teevan (Ed.), *Personal information management* (pp. 153-166). Seattle, WA: University of Washington Press.
- Soper, M. (1976). Characteristics and use of personal collections. *Library Quarterly*, 46(4), 397-415.
- Stempel, G. (1955). Increasing reliability in content analysis. *Journalism & Mass Communication Quarterly*, 32, 449-455.
- Taylor, A., & Joudrey, D. (2008). *The organization of information*. Englewood, CO: Libraries Unlimited.
- Teevan, J., Alvarado, C., Ackerman, M., & Karger, D. (2004). The perfect search engine is not enough: A study of orienteering behavior in directed search. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Austria*, 415-422.
- Teevan, J., Jones, W., & Bederson, B. (2006). Personal information management. *Communications of the ACM*, 9(1), 40-43.
- Van Someren, M., Barnard, Y., & Sandberg, J. (1994). *The think aloud method: A practical guide to modeling cognitive processes*. London, UK: Academic Press.
- Wang, P., & Soergel, D. (1998). A cognitive model of document use during a research project. Study I. document selection. *Journal of the American Society for Information Science*, 49(2), 115-133.
- Whittaker, S. (2005). Collaborative task management in email. *Human Computer Interaction*, 20, 49-88.
- Whittaker, S. (2011). Personal information management: From information consumption to curation. *Annual Review of Information Science and Technology*, 45, 3-62.
- Whittaker, S., Bellotti, V., & Gwizdka, J. (2006). Email in personal information management. *Communications of the ACM*, 49(1), 68-73.
- Whittaker, S., Bellotti, V., & Gwizdka, J. (2007). Everything through Email. In W. Jones & J. Teevan (Ed.), *Personal information management* (pp. 167-189). Seattle, WA: University of Washington Press.
- Whittaker, S., Bergman, O., & Clough, P. (2010). Easy on that trigger dad: A study of long term family photo retrieval. *Personal Ubiquitous Computing*, 14(1), 31-43.
- Whittaker, S., & Hirschberg, J. (2001). The character, value and management of personal paper archives. *ACM Transactions on Computer-Human Interaction*, 8(2), 150-170.

- Whittaker, S., Jones, Q., Nardi, B., Creech, M., Terveen, L., Isaacs., E., et al. (2004). ContactMap: Organizing communication in a social desktop. *ACM Transactions on Computer Human Interaction*, 11(4), 445-471.
- Whittaker, S., & Sidner, C. (1996). Email overload: exploring personal information management of email. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Canada*, 276-283.
- Wilson, T. (1997). Information behavior: And interdisciplinary perspective. *Information Processing and Management*, 33(4), 551-572.
- Wilson, T. (1999). Models in information behavior research. *Journal of Documentation*, 55(3), 249-270.
- Wittgenstein, L., (1953). *Philosophical investigations* (G.E.M. Anscombe, Trans.). New York, NY: Macmillan (Original work published 1945).
- Zadeh, L. (1965). Fuzzy sets, *Information and Control*, 8(3), 338-353.
- Zerubavel, E. (1991). *The fine line: Making distinctions in everyday life*. Chicago, IL: The University of Chicago Press.
- Zerubavel, E. (1996). Lumping and splitting: Notes on social classification. *Sociological Forum*, 11(3), 421-433.
- Zerubavel, E. (1997). *Social mindscapes: An invitation to cognitive sociology*. Cambridge, MA: Harvard University Press.

CURRICULUM VITAE

KYONG EUN OH

Education

- | | |
|--|------|
| Ph.D., Communication, Information and Library Studies
School of Communication & Information
Rutgers, The State University of New Jersey, New Brunswick, NJ | 2013 |
| M.A., Library and Information Science
Yonsei University, Seoul, Korea | 2008 |
| B.A., English Language and Literature & B.A., Library and Information Science
Yonsei University, Seoul, Korea | 2006 |

Publications

Journal Articles

- Oh, K., & Gwizdka, J.** (2011). Impatient opportunists: A study of technology use in a higher education classroom. *Journal of Applied Research in Higher Education*, 3(2), 81-96.
- Oh, K., & Kim, G.** (2008). A study on the convenience in finding books according to classification: Focused on the classifications in public libraries and bookstores. *Journal of Korean Society for Information Management*, 25(4), 25-42.

Book Chapter

- Oh, K.** (2010). A proposal of the topic maps for knowledge management. In W. Hastings, C. Pascal, S. Wengler, & T. Kelly (Eds.), *Knowledge work* (pp. 97-112). New Brunswick, NJ: Knowledge Institute at Rutgers University.

Conference Proceedings

- Oh, K.** (2013). The process of organizing personal information. *Advances in Classification Research: Proceedings of the 23th American Society for Information Science and Technology (ASIS&T) SIG/CR Workshop*, 5-8.
- Oh, K.** (2012). What happens once you categorize files into folders? *Proceedings of the 75th American Society for Information Science and Technology (ASIS&T) Annual Meeting, USA*, 49(1), 1-4.

- Oh, K.**, Bemis, K., Ma, X., Tremaine, M., & Silver, D. (2012). Uncovering visualization properties that confuse. *Proceedings of the 75th American Society for Information Science and Technology (ASIS&T) Annual Meeting, USA*, 49(1), 1-4.
- Oh, K.** (2012). Exploring the process of organizing personal information. *Proceedings of the 2012 iConference, Toronto, Canada*, 433-434.
- Oh, K.**, Joo, S., & Park, Y. (2012). Which facets are easy to use and useful in accessing consumer health information on the Web? A pilot study. *Proceedings of the 2012 iConference, Toronto, Canada*, 480-481.
- Oh, K.**, & Belkin, N. (2011). Cross analysis of keeping personal information in different forms. *Proceedings of the 2011 iConference, USA*, 732-733.
- Oh, K.**, Halpern, D., Chiang, J., Tremaine, M., Bemis, K., & Silver, D. (2011). Which properties in 3D visualizations cause comprehension difficulties? *Proceedings of the 74th American Society for Information Science and Technology (ASIS&T) Annual Meeting, USA*, 48(1), 1-4.
- Halpern, D., **Oh, K.**, Chiang, J., Silver, D., Tremaine, M., & Bemis, K. (2011). The paradox of visualizations making information comprehensible. *Proceedings of the IEEE VisWeek 2011, USA*, 1-2.
- Kim, G., Lee, J., & **Oh, K.** (2008). An exploratory analysis of Korean public library statistics. *Proceedings of the 15th Korean Society for Information Management, Korea*, 15, 111-116.